GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:38:47 ; Search time 8472.11 Seconds
(without alignments)
12117.322 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggcccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

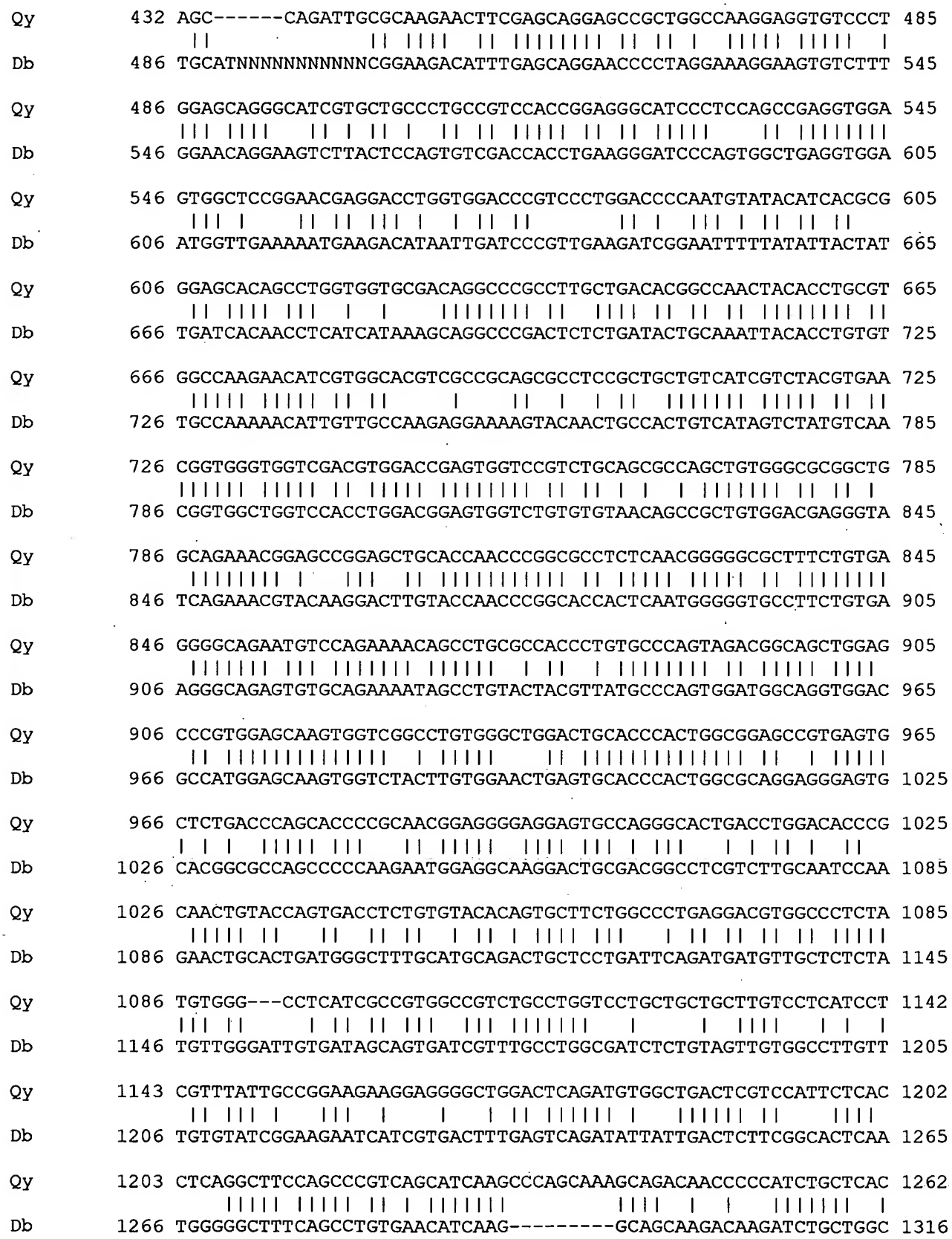
Result No.	Score	% Query Match	Length	DB	ID	Description
1	960.2	35.6	2802	9	AY406491	AY406491 Homo sapi
2	950.4	35.2	2791	9	AY406493	AY406493 Mus muscu
3	923.4	34.2	3790	3	AK031655	AK031655 Mus muscu
4	874	32.4	1852	3	CR598115	CR598115 full-leng
5	871.4	32.3	3866	3	AK018177	AK018177 Mus muscu
6	814	30.2	2802	9	AY406492	AY406492 Pan trogl
7	810.4	30.0	2532	9	AY411747	AY411747 Homo sapi
8	780.4	28.9	2532	9	AY411749	AY411749 Mus muscu

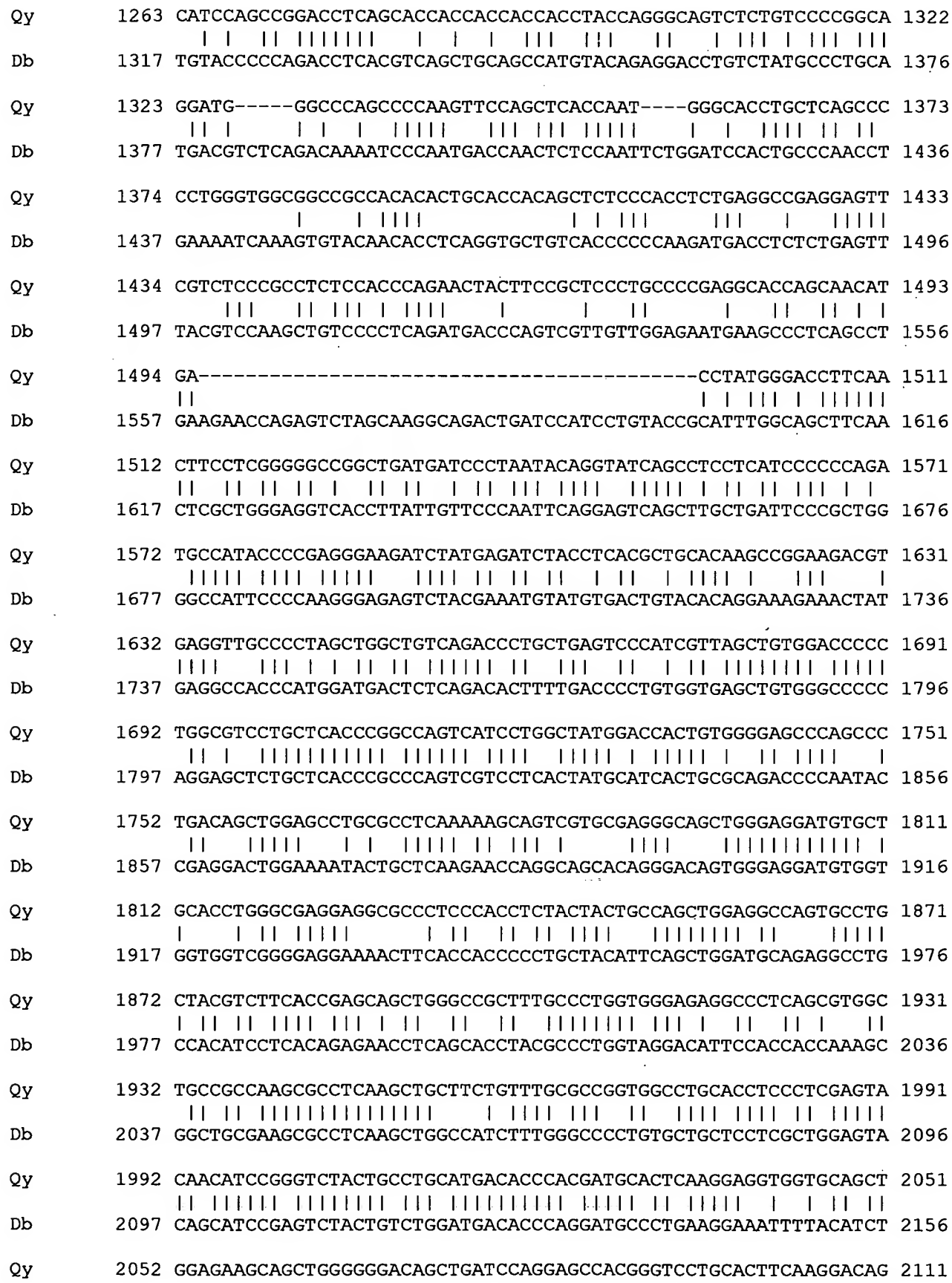
9	768.6	28.5	1034	4	BI758231	BI758231	603029876
10	736.6	27.3	1532	3	BC033727	BC033727	Homo sapi
11	735.2	27.3	788	1	AI951556	AI951556	wv36f04.x
c 12	721.6	26.8	796	5	BX348193	BX348193	BX348193
13	713.4	26.5	818	4	BI818609	BI818609	603033362
14	678.2	25.1	2532	9	AY411748	AY411748	Pan trogl
15	672.6	24.9	934	2	BF311804	BF311804	601897316
16	665	24.7	859	2	BF311896	BF311896	601897733
17	647.4	24.0	853	5	BX364574	BX364574	BX364574
18	621	23.0	900	5	BX345406	BX345406	BX345406
19	618.8	22.9	756	5	BU612387	BU612387	UI-M-EW0-
20	610.2	22.6	2775	9	AY401471	AY401471	Mus muscu
21	603.2	22.4	874	5	BQ689148	BQ689148	AGENCOURT
22	600.6	22.3	889	5	BQ691915	BQ691915	AGENCOURT
23	600.2	22.3	977	5	BX345407	BX345407	BX345407
24	599	22.2	2775	9	AY401469	AY401469	Homo sapi
25	595.2	22.1	601	1	AL516580	AL516580	AL516580
26	589.4	21.9	604	2	BE314370	BE314370	601147261
27	586.6	21.8	1072	5	BX422753	BX422753	BX422753
28	566.6	21.0	678	6	CA749784	CA749784	UI-M-FD0-
29	561	20.8	1175	2	BF530640	BF530640	602071931
30	555.6	20.6	788	6	CA317532	CA317532	UI-M-FW0-
31	555	20.6	572	7	CR554569	CR554569	DKFZp459I
32	491.2	18.2	499	7	CR747398	CR747398	CR747398
33	486.2	18.0	548	5	BX452510	BX452510	BX452510
34	471.2	17.5	540	4	BI467799	BI467799	389583 MA
35	470.8	17.5	2507	9	AY401470	AY401470	Pan trogl
36	468.6	17.4	824	4	BI737024	BI737024	603360874
37	438.2	16.2	675	6	CA315487	CA315487	UI-M-FW0-
38	433	16.1	471	5	BX282095	BX282095	BX282095
39	428.4	15.9	1147	5	BU840446	BU840446	AGENCOURT
40	424	15.7	843	4	BG913440	BG913440	602811321
41	417.8	15.5	856	7	CN164143	CN164143	994266 MA
42	411.8	15.3	749	7	CF735417	CF735417	UI-M-HB0-
43	410.6	15.2	1111	4	BG298307	BG298307	602397080
44	409.8	15.2	751	7	CF735550	CF735550	UI-M-HB0-
45	408.8	15.2	460	2	BF443156	BF443156	260542 MA

ALIGNMENTS

RESULT 1
AY406491
LOCUS AY406491 2802 bp DNA linear GSS 15-DEC-2003
DEFINITION Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY406491
VERSION AY406491.1 GI:39762465
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2802)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

```
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2802)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2802
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             35.6%;  Score 960.2;  DB 9;  Length 2802;
  Best Local Similarity   61.3%;  Pred. No. 1.9e-194;
  Matches 1682;  Conservative    0;  Mismatches  995;  Indels   69;  Gaps    6;

Qy         12 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 71
              || ||  ||     |   |  ||||  ||  | |  || | || |   |  || |   |
Db         66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125

Qy         72 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 131
              ||   |      |       |  |  |  |  | ||| ||     || || |  || ||
Db        126 TGATGACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCC 185

Qy        132 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 191
               || |||||  | |||||| || || |  || || || ||||| ||||| |||   ||
Db        186 ACATTTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

Qy        192 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA 251
               || || ||    || ||||| |||||||  |||||||| ||  | || |||||  |  ||
Db        246 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA 305

Qy        252 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGT 311
              |  ||||||| |  | ||    ||   ||| |  |     || ||      |  ||| ||
Db        306 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGT 365

Qy        312 CCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTG 371
               |||||  |  | ||  | ||||| || ||  |  | || || |  || || ||||||||
Db        366 GAGCATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG 425

Qy        372 CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT 431
              |||||| ||||| ||||||||| |||| ||||| ||||| | |||||| ||  | |||||
Db        426 CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT 485
```

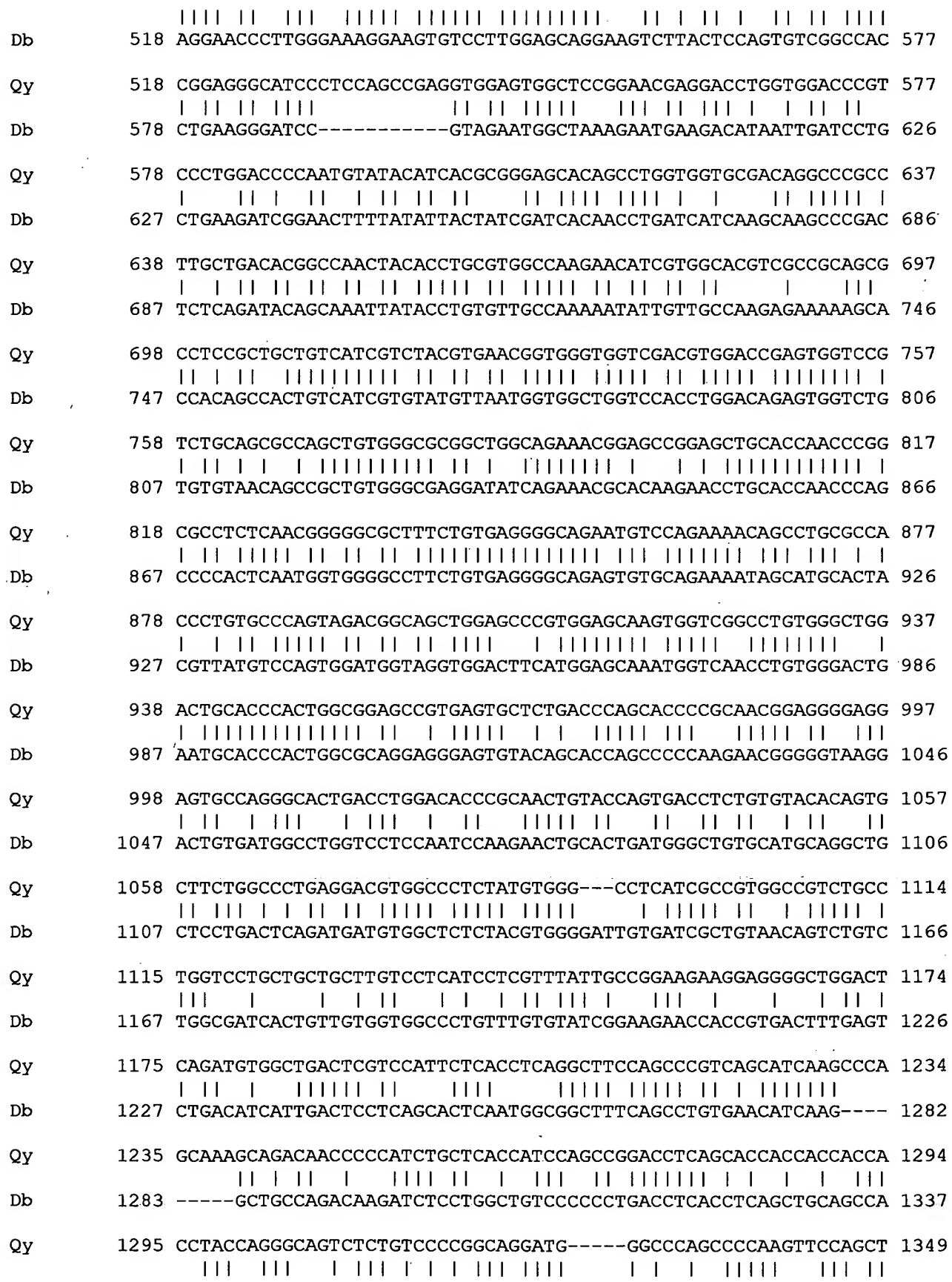


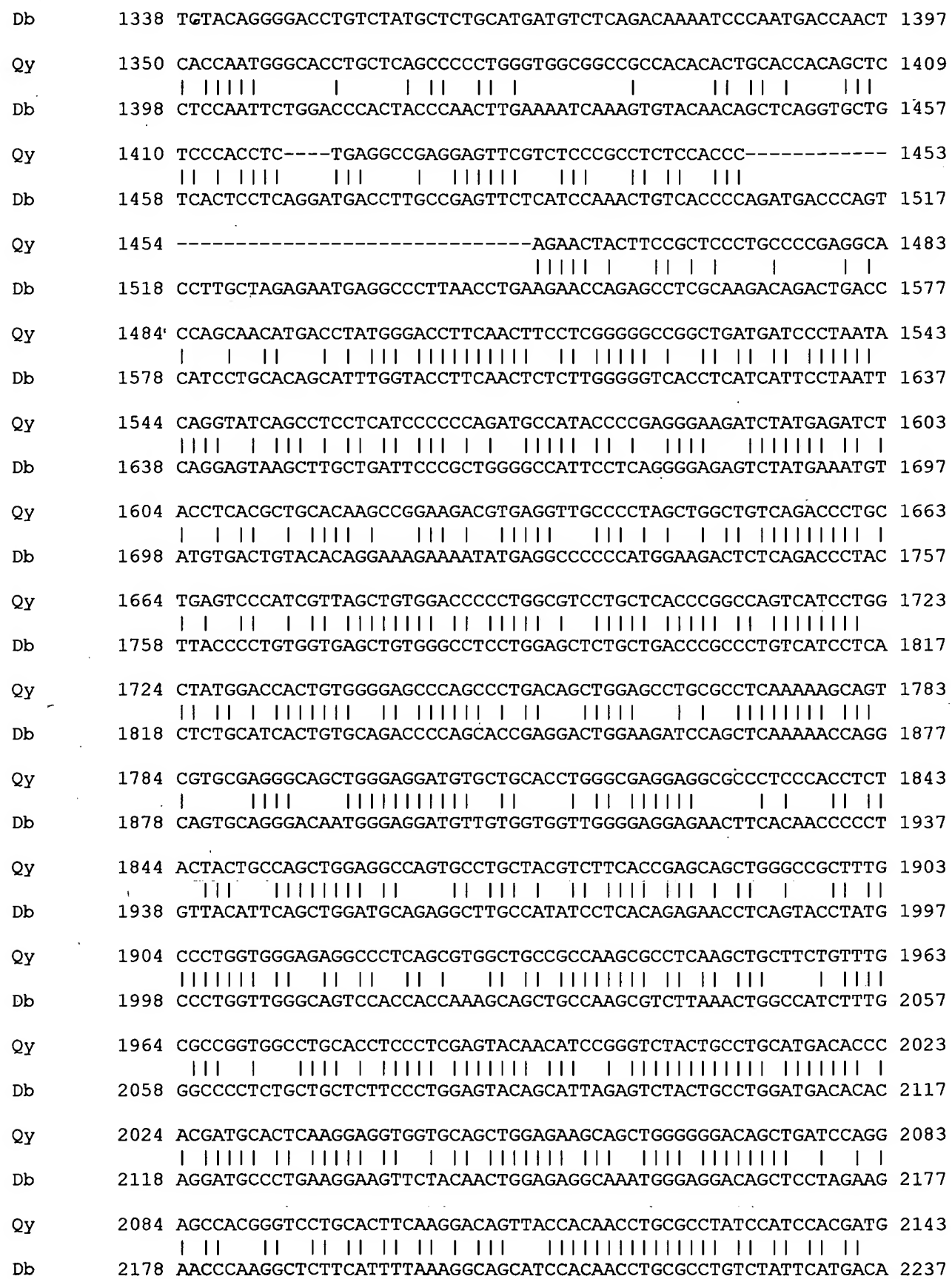


```
                |||| ||| |||| |||||||| | | || ||   ||  || || || || | |||
Db       2157 TGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAG 2216

Qy       2112 TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 2171
                ||||||||||||||| || || |||||| |  ||   ||||| |||||||| ||  |
Db       2217 CACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCATTCCCTCTGGAAGAGCAAATT 2276

Qy       2172 CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2231
                || |  | || ||||| || || ||||| ||  | |||| |||  | ||  |  || |
Db       2277 GCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCT 2336

Qy       2232 GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2291
                |||||||||||||||| |||||  |  | ||||  | |||    || ||||  ||||| ||
Db       2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396

Qy       2292 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351
                || ||| |||||||||| || || ||||||| ||||    |||||| || | |  |  |
Db       2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456

Qy       2352 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411
                |||  | |        ||  ||| |   |||||   ||         || |  |  |
Db       2457 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516

Qy       2412 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471
                ||| |||||||| ||||  |||||  |||  || |||||||||  |   |  |||||||||
Db       2517 GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576

Qy       2472 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531
                | |||   | | | ||    |||||| |||  |||||||| || ||  ||||||||||
Db       2577 TGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTGGACAG 2636

Qy       2532 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2591
                |  | |  | |||||||| ||||  ||||||| || | | | |||||  | || |||||
Db       2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696

Qy       2592 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2651
                || | | |||||||  | || |||||.|||  ||||||  || |   |||  | | ||||
Db       2697 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756

Qy       2652 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
                   |  ||  | |   | | |  ||   | || |   |||  | |
Db       2757 AAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
```

```
RESULT 2
AY406493
LOCUS       AY406493               2791 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY406493
VERSION     AY406493.1  GI:39762467
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2791)
```

```
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2791)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2791
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>2791
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             35.2%;  Score 950.4;  DB 9;  Length 2791;
  Best Local Similarity   62.3%;  Pred. No. 2.3e-192;
  Matches 1653;  Conservative    0;  Mismatches  921;  Indels   80;  Gaps    7;

Qy        104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
              | ||   || |   || || |   | || ||||||||  | ||||||||||| |  ||||
Db        158 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 217

Qy        164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223
              | || ||||||||||| |||   ||    || || |||   || ||||| ||||||| ||
Db        218 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 277

Qy        224 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 283
              |||||||||||  | ||||||||  |  |||  |||||||||  | ||    ||   ||| |
Db        278 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 337

Qy        284 GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 343
                | |  ||| ||        |  || ||  ||||| |  | ||| | ||||||||  |||
Db        338 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 397

Qy        344 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 403
              |  |||| ||||  || || |||||||||||||| ||||| ||||||||  | ||||| |
Db        398 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 457

Qy        404 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAG------ATTGCGCAAGAACTTCGAGC 457
              | ||||||| |||||| ||| | ||||| ||           |||| ||||  |||||||
Db        458 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTTTCCAGATCTGCGGAAGACATTCGAGC 517

Qy        458 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 517
```





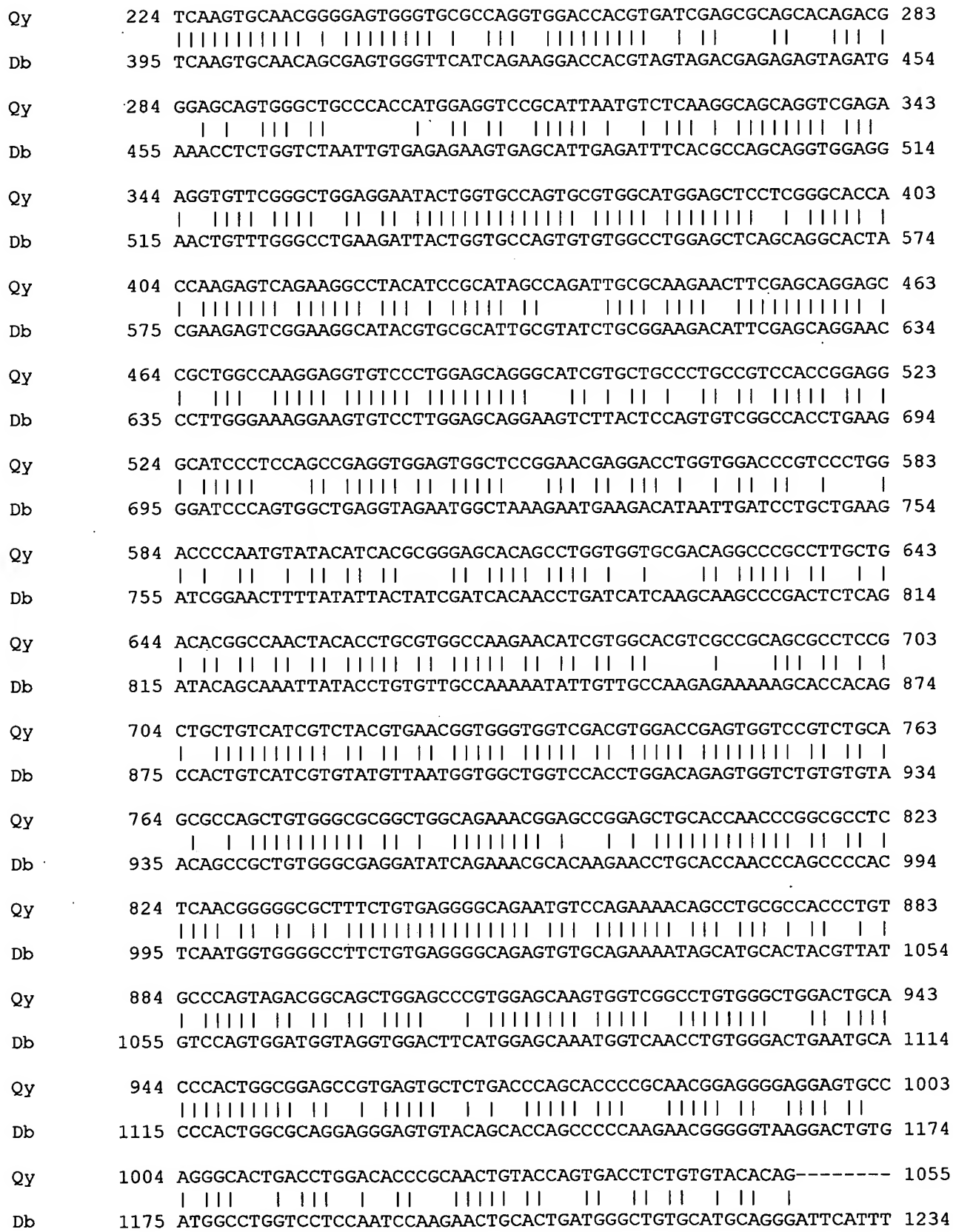
```
Qy       2144 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATC 2203
              |  ||   ||||| |||||||| ||  | || |  |  || ||||| || || ||||| |
Db       2238 TCGCCCATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACC 2297

Qy       2204 ACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCC 2263
              ||||||||| |||| | ||  |  || | |||||||||||||| |||||  |  |||||| 
Db       2298 ACATCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCC 2357

Qy       2264 CCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGA 2323
                | |||    || ||||  ||||| || || ||| ||||||| || || || |||||||
Db       2358 TAAACACAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGA 2417

Qy       2324 GCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGG 2383
              ||||    ||||||  |   |  | ||||  | |         ||  | |     ||||
Db       2418 TCTTCCAGCTCAACTGTACTGTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGG 2477

Qy       2384 AGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCA 2443
              |   ||          || |  |  |    || || ||||| ||||  || ||  |||  |
Db       2478 ACCCTGCTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTA 2537

Qy       2444 TTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGA 2503
              | ||||||||| ||     |||||||||  | ||   |   | ||    |||||| |||
Db       2538 TCCGGCAGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGA 2597

Qy       2504 CTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCC 2563
                |||||||| ||||||  |||||||||||  |  | |  | |||||||  ||||  | ||||
Db       2598 TGCTGGCCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCC 2657

Qy       2564 CCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCC 2623
              | ||  | | | |||||  | || |||||  || | | |||||||  | || ||||| |||
Db       2658 CAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCA 2717

Qy       2624 AGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGG 2683
              ||||||  ||  |   |||  | | ||||    |  ||  | |   | | |   |   | |
Db       2718 TGCTGGCAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCCTTGGCAGCAG 2777

Qy       2684 AGGCTGAGTGCTGA 2697
              | |   |||  |||
Db       2778 AAGGACAGTATTGA 2791
```

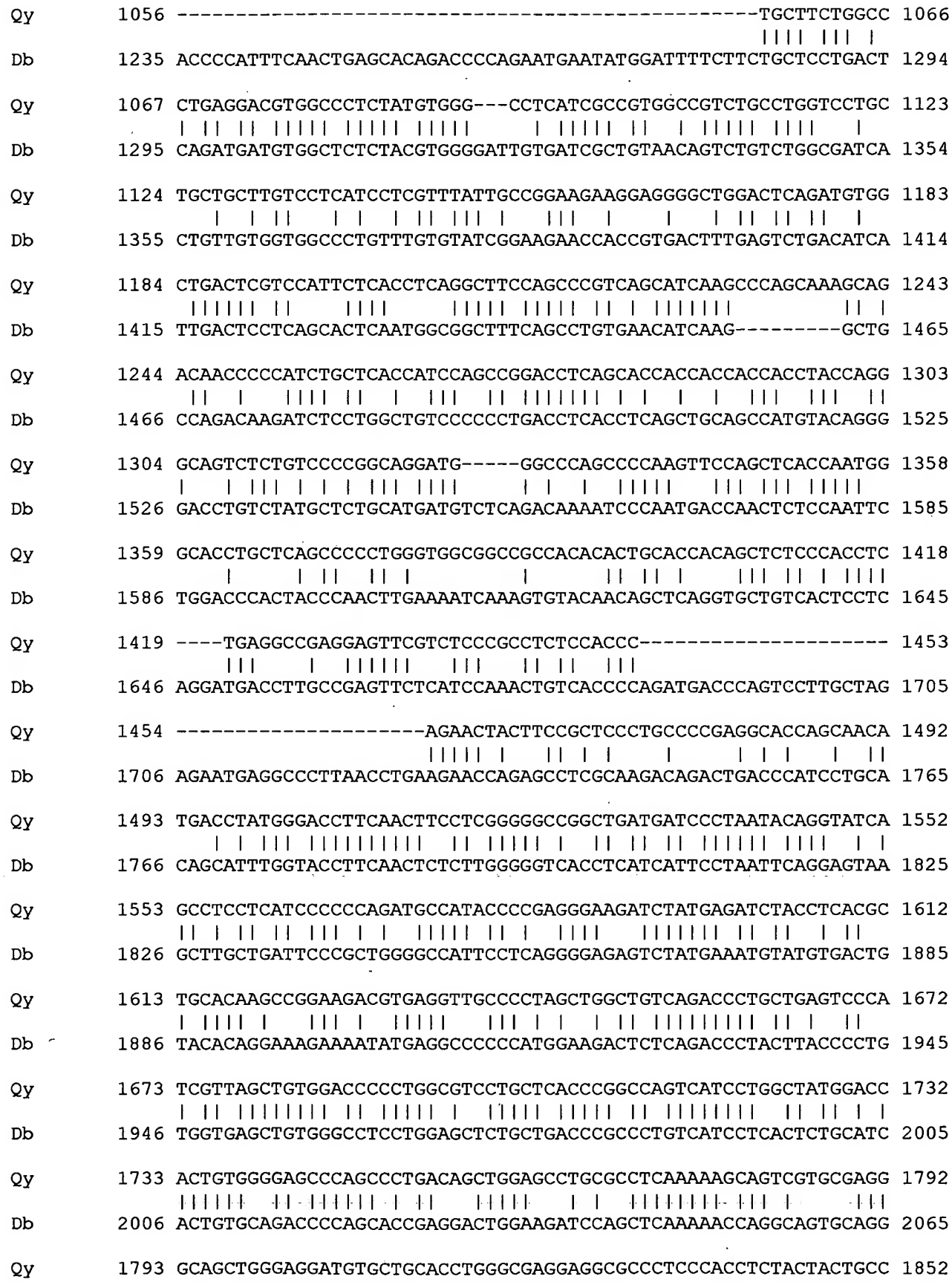
RESULT 3
AK031655

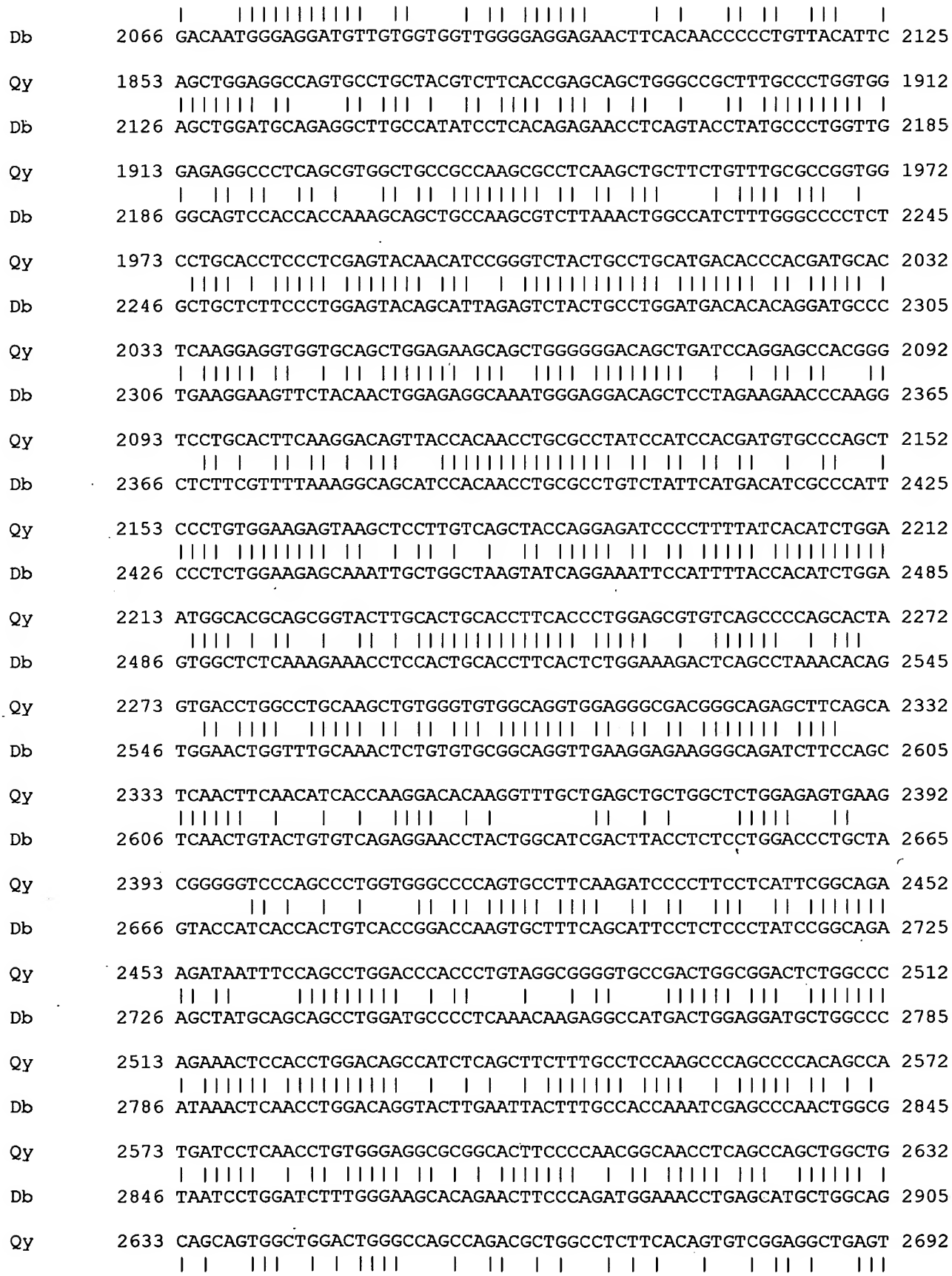
LOCUS	AK031655 3790 bp mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030473H24 product:unc5 homolog (C. elegans) 3, full insert sequence.
ACCESSION	AK031655
VERSION	AK031655.1 GI:26327502
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3790)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

```
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .3790
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:6030473H24"
                     /db_xref="taxon:10090"
                     /clone="6030473H24"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="13 days embryo"
     CDS             118. .2970
                     /note="unnamed protein product; putative
                     unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
                     GB|NM_009472, evidence: BLASTN, 99%, match=464)"
                     /codon_start=1
                     /protein_id="BAC27495.1"
                     /db_xref="GI:26327503"
                     /translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD
                     SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA
                     RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT
                     PQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPN
                     SGVSLLIPAGAIPQGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPV
                     ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTEN
                     LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
                     GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHC
                     TFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVT
                     GPSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDL
                     WEAQNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQY"
ORIGIN

  Query Match             34.2%;  Score 923.4;  DB 3;  Length 3790;
  Best Local Similarity   61.3%;  Pred. No. 1.4e-186;
  Matches 1659;  Conservative    0;  Mismatches  926;  Indels  120;  Gaps    6;

Qy        104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
              | ||   || |   || || |   | || ||||||||  | ||||||||||| |  ||||
Db        275 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 334

Qy        164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223
              | || ||||||||||| |||   ||    || || |||   || ||||| ||||||| ||
Db        335 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 394
```







```
Db       2906 CCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTACTTGGCAGCAGAAGGACAGT 2965

Qy       2693 GCTGA 2697
                 |||
Db       2966 ATTGA 2970


RESULT 4
CR598115
LOCUS       CR598115               1852 bp    mRNA    linear   HTC 21-JUL-2004
DEFINITION  full-length cDNA clone CS0DA006YG16 of Neuroblastoma of Homo
            sapiens (human).
ACCESSION   CR598115
VERSION     CR598115.1  GI:50478922
KEYWORDS    HTC; CNSLT_cDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1852)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
  REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue
REFERENCE   2  (bases 1 to 1852)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1852
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DA006YG16"
                     /tissue_type="Neuroblastoma"
                     /plasmid="pCMVSPORT_6"
ORIGIN

  Query Match             32.4%;  Score 874;  DB 3;  Length 1852;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-176;
  Matches  874;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1824 GGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 1883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 60

Qy       1884 CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 1943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 120

Qy       1944 CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGT 2003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGT 180

Qy       2004 CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT 2063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT 240

Qy       2064 GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT 2123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT 300

Qy       2124 GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA 2183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA 360

Qy       2184 CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT 2243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT 420

Qy       2244 CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA 2303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA 480

Qy       2304 GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT 2363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT 540

Qy       2364 TGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC 2423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC 600

Qy       2424 CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAG 2483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAG 660

Qy       2484 GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT 2543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT 720

Qy       2544 CTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTT 2603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTT 780

Qy       2604 CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGC 2663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGC 840

Qy       2664 TGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
              |||||||||||||||||||||||||||||||||
Db        841 TGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 874
```

RESULT 5
AK018177

LOCUS AK018177 3866 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male medulla oblongata cDNA, RIKEN full-length enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION AK018177
VERSION AK018177.1 GI:12857775
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3866)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,

Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES Location/Qualifiers

source 1. .3866
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6330415E02"
/db_xref="taxon:10090"
/clone="6330415E02"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

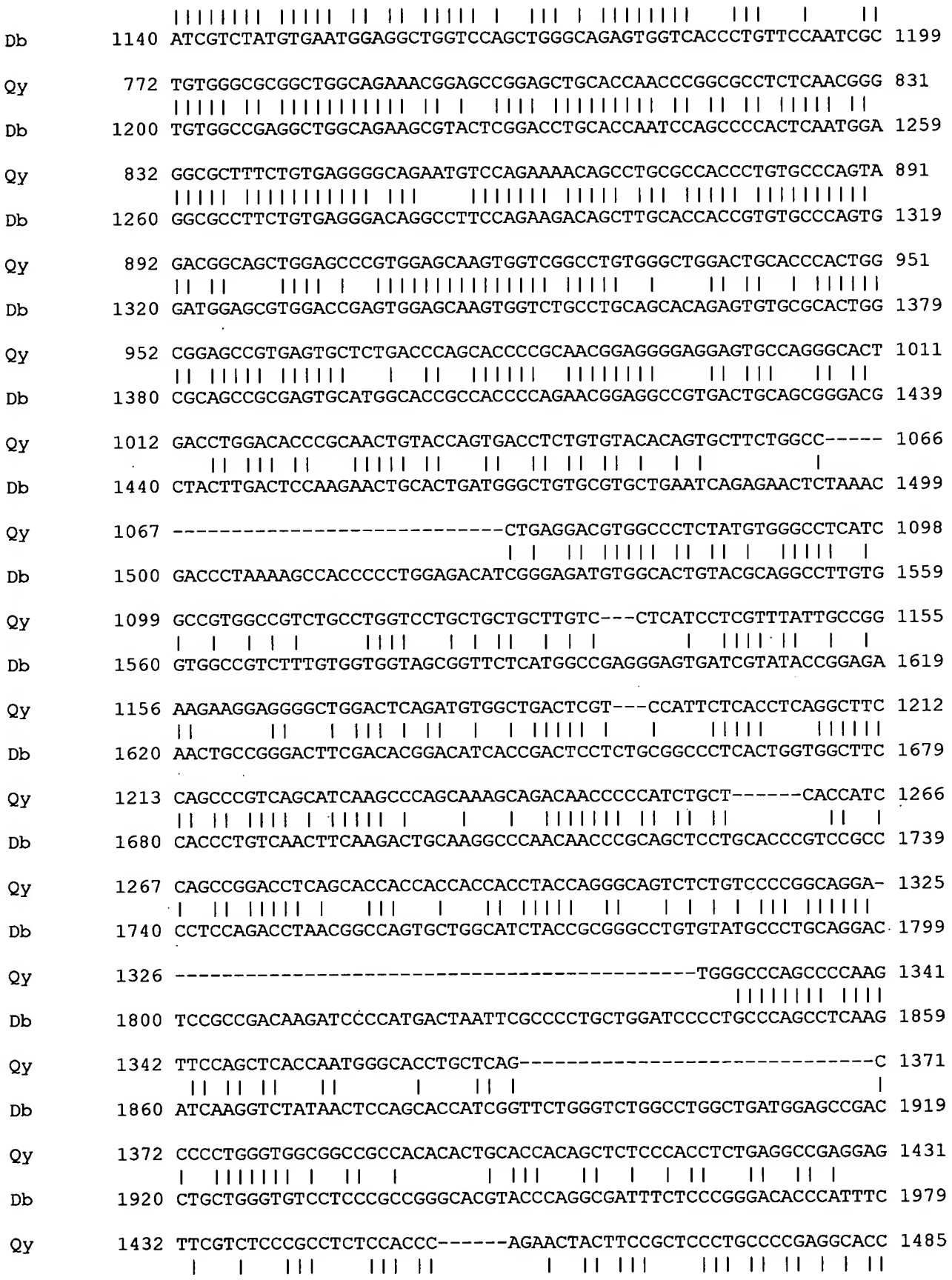
CDS 417. .3254
/note="unnamed protein product; TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus] (SPTR|O08722, evidence: FASTY, 96.5%ID, 100%length, match=2835)
putative"
/codon_start=1
/protein_id="BAB31108.1"
/db_xref="GI:12857776"
/translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPA
EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG
LRVREVQIEVSRQQVEELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPL
AKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS
DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNP
APLNGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNG
GRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM

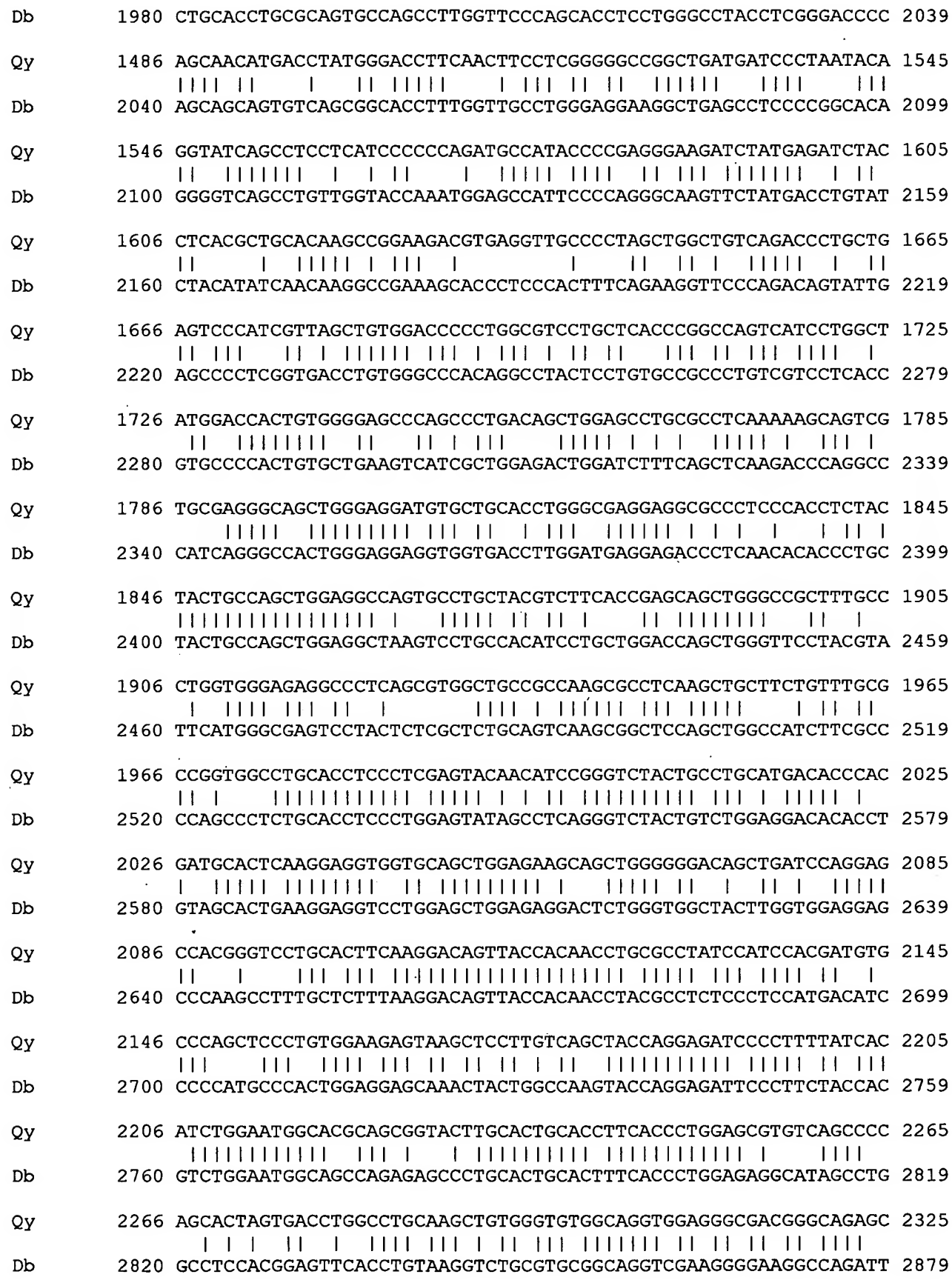
AEGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAG
IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP
GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLL
VPNGAIPQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPH
CAEVIAGDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF
MGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE
EPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER
HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGNAITTQLGPYAF
KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ
QDDGDLNSLASALEEMGKSEMLVAMATDGDC"

ORIGIN

Query Match 32.3%; Score 871.4; DB 3; Length 3866;
Best Local Similarity 60.6%; Pred. No. 1.9e-175;
Matches 1645; Conservative 0; Mismatches 941; Indels 129; Gaps 8;

```
Qy        112 GCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAG 171
              |||    |||| | ||| ||  |||||||  |||||||  |||| |  |||||||| |||
Db        540 GCCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG 599

Qy        172 AACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGC 231
              ||||||||||||   ||    ||||  ||| | || ||||||||||||| |||||||||
Db        600 AACAAGCCAGTGGAACTGCACTGCAGAGCCTTCCCTGCCACGCAGATCTACTTCAAGTGT 659

Qy        232 AACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT 291
              || || ||||||||  |||||   ||||||||  |   ||   |||   || | |  ||
Db        660 AATGGCGAGTGGGTCAGCCAGAATGACCACGTCACACAGGAGAGCCTGGATGAGGCCACA 719

Qy        292 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 351
              ||  |||        |||||  |  ||  | || ||| ||||||| || ||| |  | |||
Db        720 GGCTTGCGGGTGCGAGAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTC 779

Qy        352 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 411
              ||||| ||||| ||||||||||||||||||| |||||||| ||||| || |||||||||
Db        780 GGGCTCGAGGACTACTGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGT 839

Qy        412 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 471
              |     |||||||||||||| |||   |||||||||||||| || |||||||| ||||||
Db        840 CGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCC 899

Qy        472 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 531
              ||||||||  || |||| || |  || | ||||  ||||| ||||||||||  | |||
Db        900 AAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCT 959

Qy        532 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 591
                 || |||||||| ||||||  ||| ||  ||  |  | |||||  | | |||| | ||
Db        960 GTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAAC 1019

Qy        592 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 651
               |     ||||    || |||| |||  |  | || |||||  |||||  | |||||||||
Db       1020 TTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCC 1079

Qy        652 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 711
              ||||||||||| ||||||||||  ||||||||    |||||  ||| ||  | ||  | |||
Db       1080 AACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCCGGAGCACCGCGGCCACAGTC 1139

Qy        712 ATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC 771
```

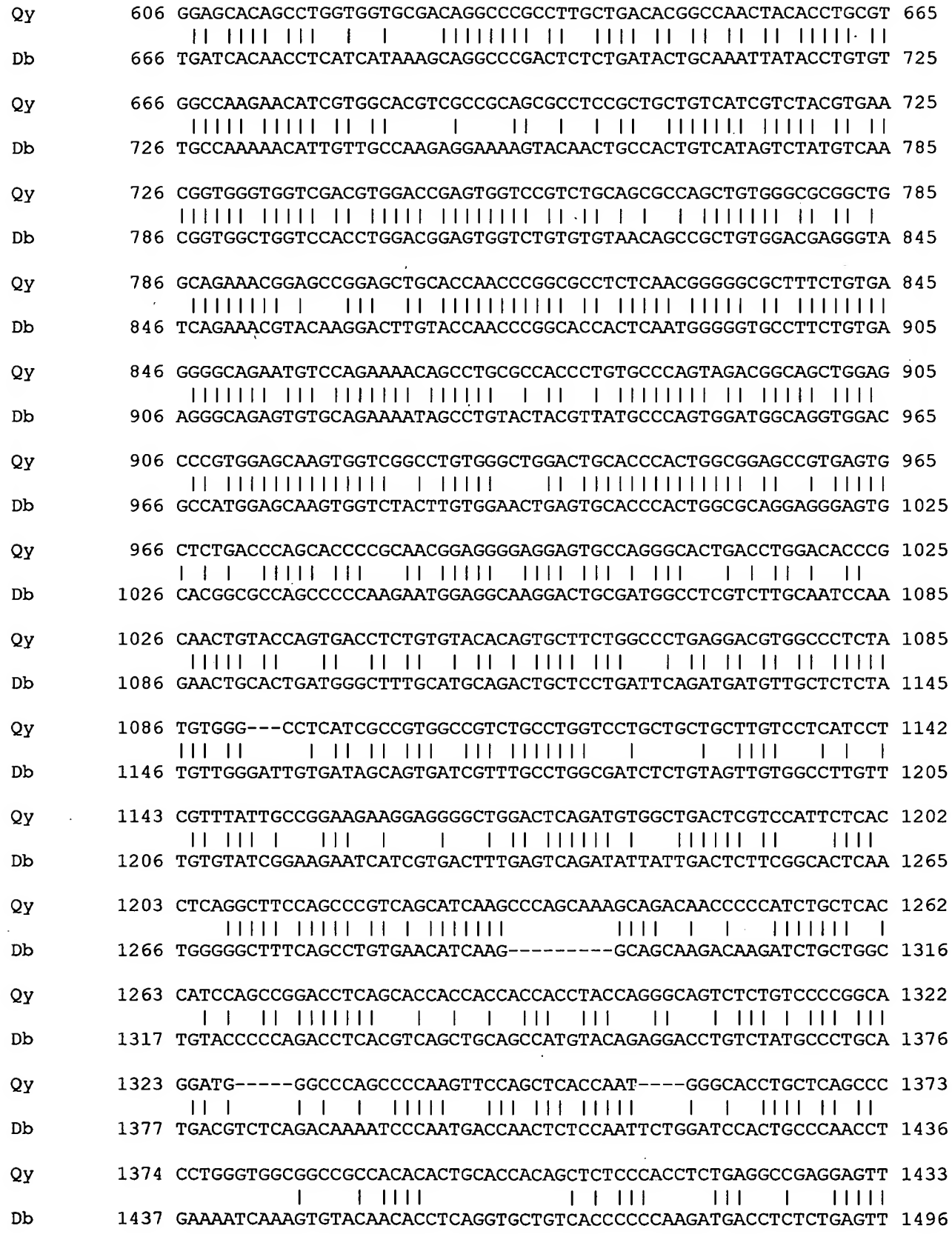


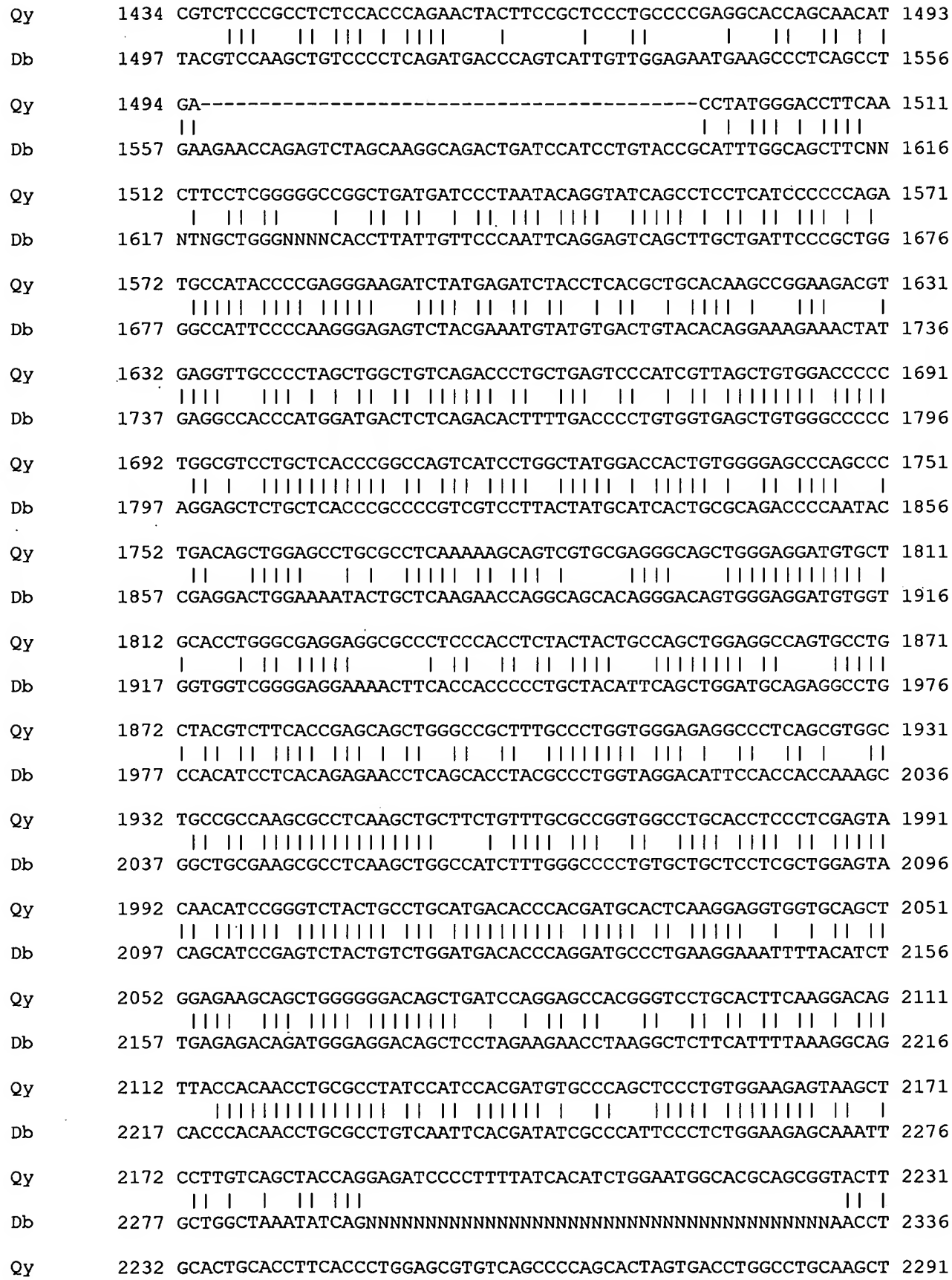


```
Qy       2326 TTCAGCATCAACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTG 2382
              |||    |  ||   |   |  ||   | |    |  | |     ||    ||
Db       2880 TTCCAGCTGCACACAACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCT 2939

Qy       2383 GAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTC 2442
              |     |     |   || |  ||| | |||| |||  |||||||||||| ||| |   |
Db       2940 GCCCCGGGCAATGCCATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCC 2999

Qy       2443 ATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGG 2502
              ||| || || |||||      |||||||||| | |||      |||||   ||||||| ||
Db       3000 ATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCCGACTCCCGGGGCAACGACTGGAGG 3059

Qy       2503 ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGC 2562
                  |||| |||||  ||   | |||||  |  | || | || ||| ||| ||||  | ||
Db       3060 CTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGT 3119

Qy       2563 CCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGC 2622
              |||||||   | |||  |  |||| |||||  || |||||        ||||  |||||| |
Db       3120 CCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAAC 3179

Qy       2623 CAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCG 2682
                 |||||    ||  |||  |   ||||| ||   ||        |   ||  |   |
Db       3180 AGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACA 3239

Qy       2683 GAGGCTGAGTGCTGA 2697
              ||  |  || ||||||
Db       3240 GATGGCGATTGCTGA 3254
```

RESULT 6
AY406492
LOCUS AY406492 2802 bp DNA linear GSS 15-DEC-2003
DEFINITION Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY406492
VERSION AY406492.1 GI:39762466
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 2802)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2802)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission

```
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2802
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>2802
                     /gene="UNC5C"
                     /locus_tag="HCM2575"
ORIGIN

  Query Match             30.2%;  Score 814;  DB 9;  Length 2802;
  Best Local Similarity   53.8%;  Pred. No. 3.4e-163;
  Matches 1477;  Conservative    0;  Mismatches 1200;  Indels   69;  Gaps    6;

Qy          12 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 71
               || ||  ||      |   |  ||||  ||  | | || | || |   |  || |   |
Db          66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125

Qy          72 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 131
               ||   |       |        |  |  |  |  | |||   |      || || |  || ||
Db         126 TGATGACTTTTTTCATGAACTCCCAGAAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC 185

Qy         132 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 191
                || |||||  | ||  || |||||  |  || || || ||||| ||||| |||   ||
Db         186 ACATTTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245

Qy         192 GTGCAAGGCC------GTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGT 245
                || || ||       |
Db         246 CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 305

Qy         246 GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCAT 305

Db         306 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 365

Qy         306 GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATA 365

Db         366 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 425

Qy         366 CTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACAT 425

Db         426 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 485

Qy         426 CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 485
                           || ||||  || |||||||| || || |  ||||| |||||  |
Db         486 NNNNNNNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAAGTGTCTTT 545

Qy         486 GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 545
               ||| ||||   || | || |  || || ||||| || || |||||    || |||
Db         546 GGAACAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605

Qy         546 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 605
                     || || ||| |  | || ||       || |  ||| | || || ||
Db         606 NNNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665
```





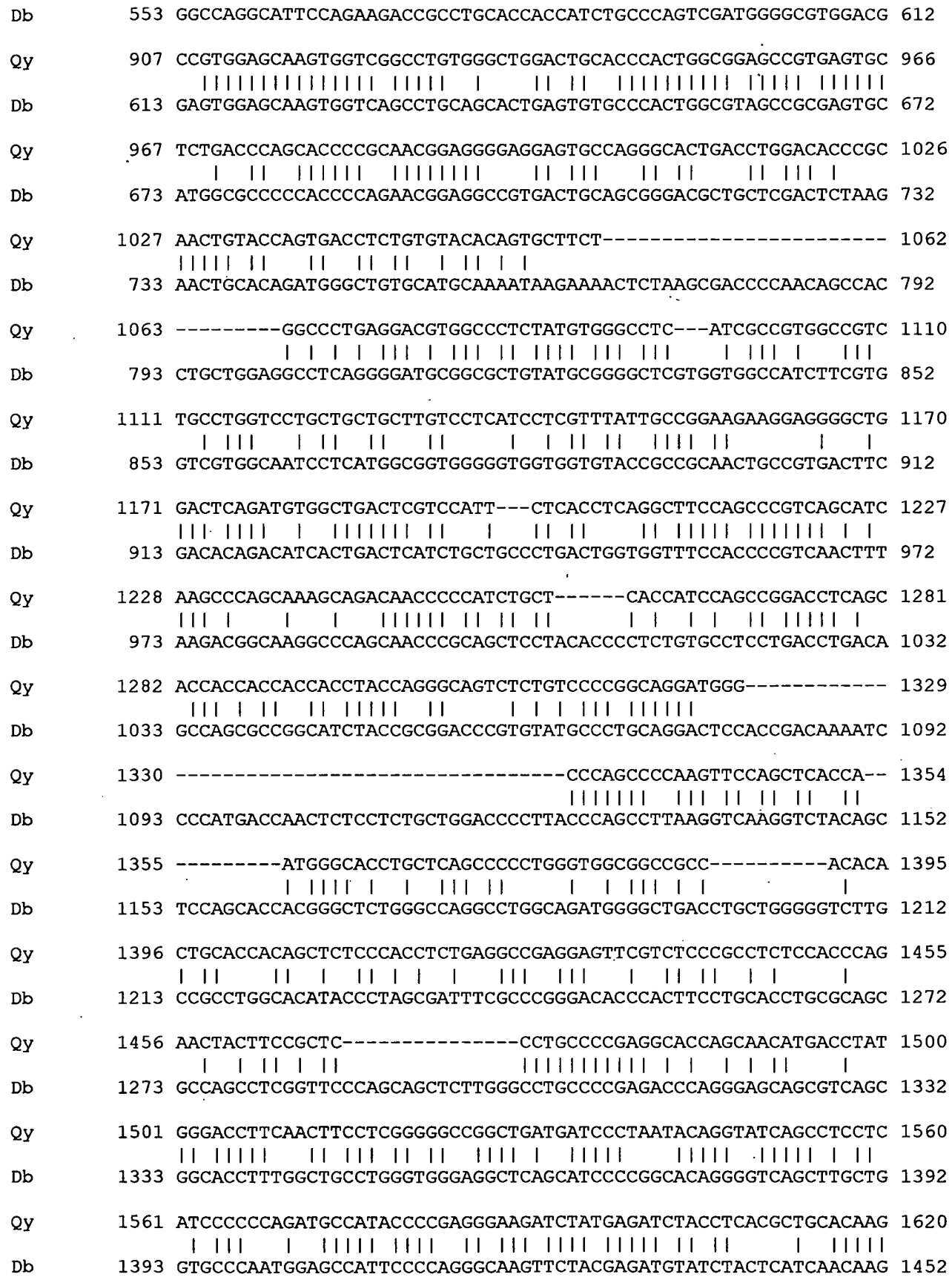
```
Db       2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396

Qy       2292 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351

Db       2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456

Qy       2352 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411

Db       2457 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516

Qy       2412 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471

Db       2517 GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576

Qy       2472 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531

Db       2577 TGCCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCNNNNNNNNNNNNN 2636

Qy       2532 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2591

Db       2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696

Qy       2592 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2651

Db       2697 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756

Qy       2652 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697

Db       2757 AAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
```

RESULT 7
AY411747
LOCUS AY411747 2532 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY411747
VERSION AY411747.1 GI:39767715
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2532)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

```
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>2532
                     /locus_tag="HCM4327"
ORIGIN
```

Query Match 30.0%; Score 810.4; DB 9; Length 2532;
Best Local Similarity 61.1%; Pred. No. 2e-162;
Matches 1540; Conservative 0; Mismatches 851; Indels 129; Gaps 9;

```
Qy        307 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 366
              ||||| |  ||  | || ||  |||||||||| ||| || | || |||||||||| |||
Db         13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy        367 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 426
              |||||||||||||||||| |||||||| ||||||||||||||||||     |||||| ||
Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCCGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy        427 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 486
              ||||| |||    |||||||||||||||| |||||||| |||| |||||||||| |||||
Db        133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTG 192

Qy        487 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 546
              || || |   |  | ||||  ||||| || ||||||||  | |||   ||||||||||
Db        193 GACCATGAGGTTCTCCTGCAGTGCCGCCCGCCGGAGGGGGTGCCTGTGGCCGAGGTGGAA 252

Qy        547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 606
              ||||||  ||| |||||  |  | |||||  ||| |||| ||||  |     ||||
Db        253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy        607 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 666
              || |||| |||  |  | || ||||||||||  | ||||| |||||||| |||||||||
Db        313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372

Qy        667 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 726
              ||||||||||||||||    ||||| ||| || | ||  | ||||||||||||||||||
Db        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432

Qy        727 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 786
              || || ||||| |  ||| | ||||||||   |||| |  | ||||||| || ||||||
Db        433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492

Qy        787 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 846
              ||||| || | ||||| |||||||||||| || || |||||||| || || ||||| |||
Db        493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552

Qy        847 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC 906
              || |||    ||||||| || |||||| ||||| | |||||||| || ||    ||||
```



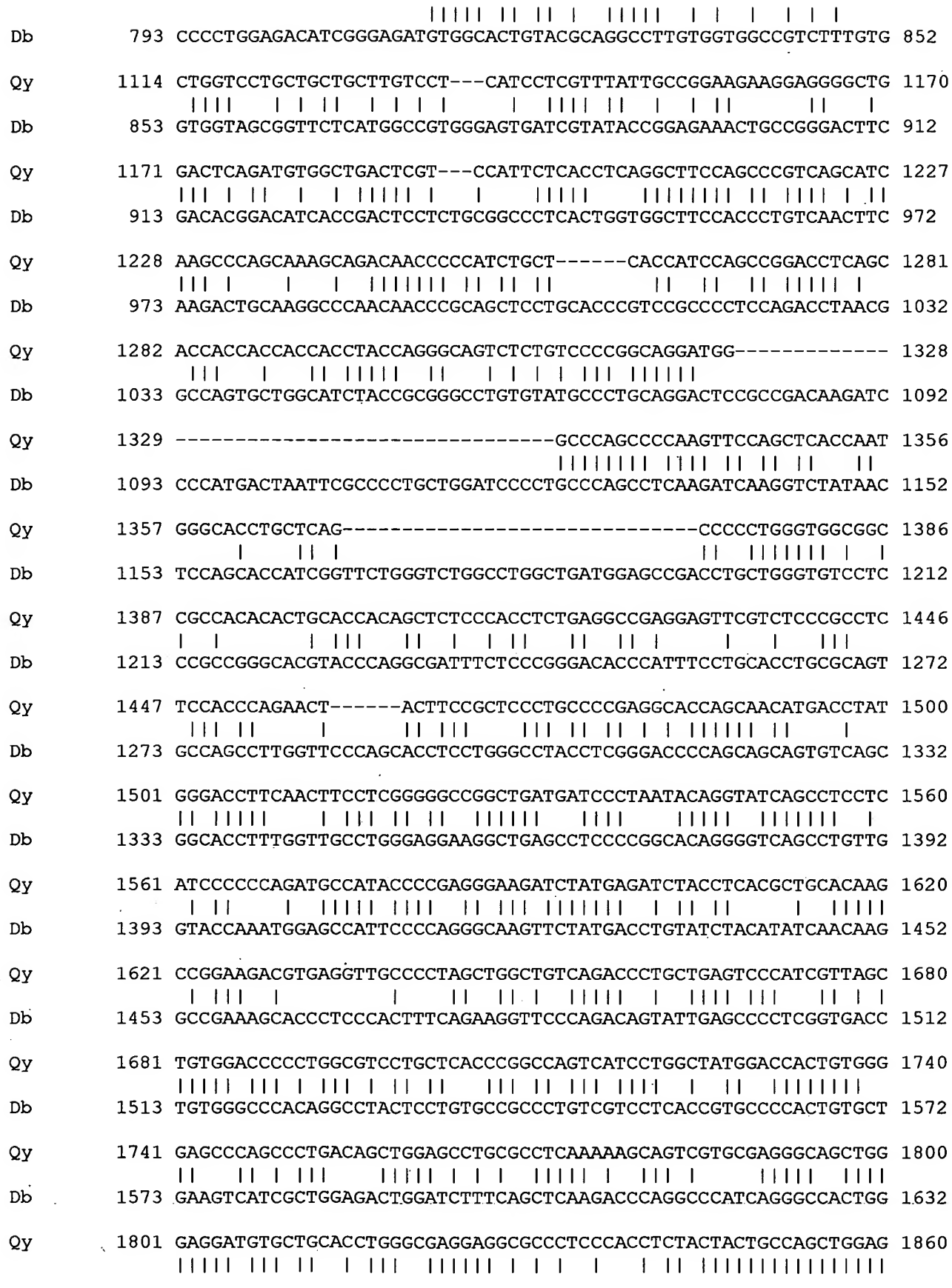
```
Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
Db       1453 GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACC 1512

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
Db       1513 TGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCC 1572

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
Db       1573 GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG 1632

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
Db       1633 GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 1692

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
Db       1693 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 1752

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
Db       1753 TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC 1812

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
Db       1813 TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG 1872

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
Db       1873 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 1932

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
Db       1933 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 1992

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
Db       1993 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC 2052

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Db       2053 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2112

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Db       2113 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2172

Qy       2341 AACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2397
Db       2173 ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT 2232

Qy       2398 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2457
Db       2233 GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA 2292
```

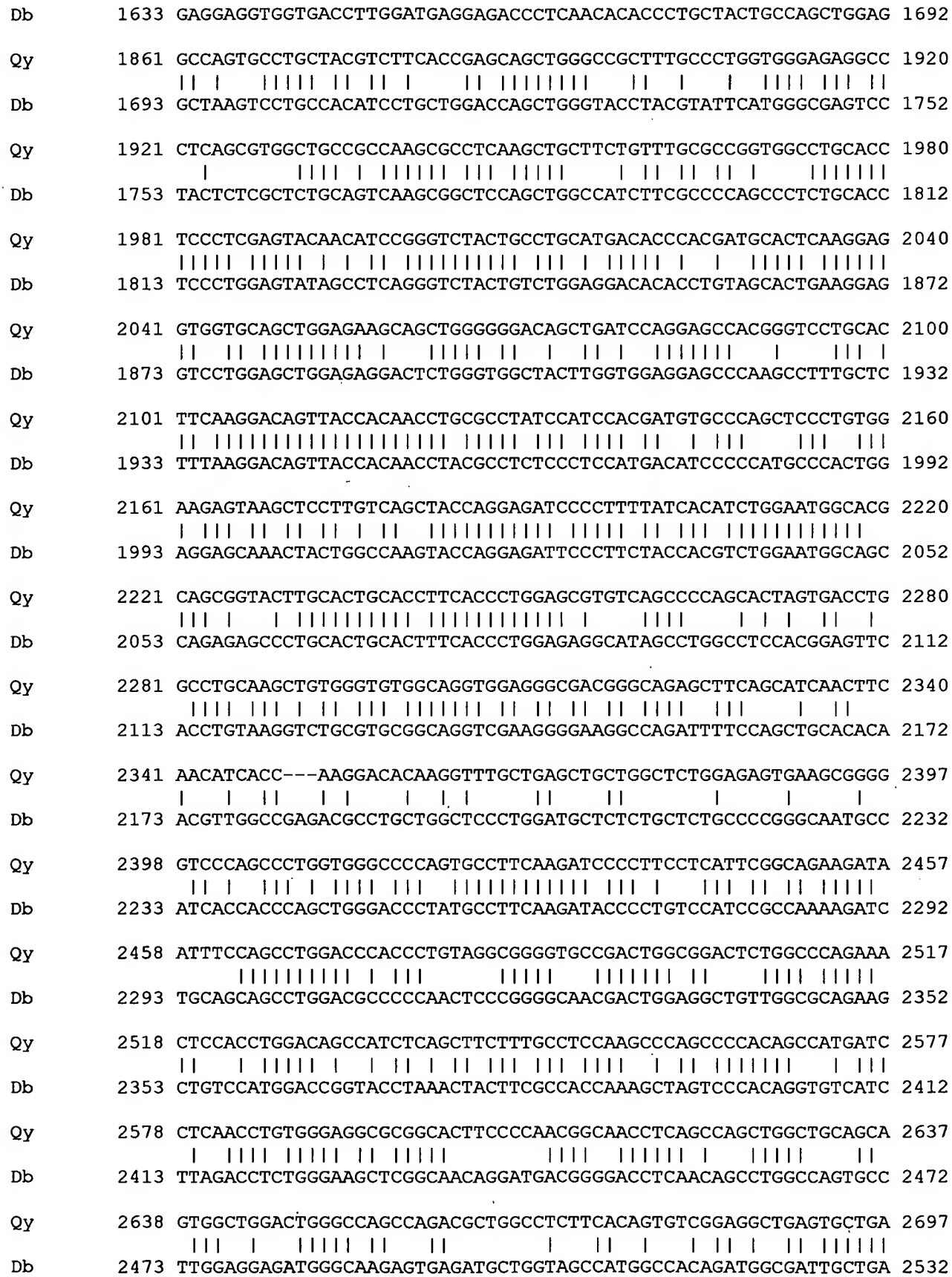
```
Qy       2458 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2517
                  ||||||  ||   | |||       |||||    ||||||||||    | ||  |||||
Db       2293 TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG 2352

Qy       2518 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2577
              |||    |||||  |   | ||  |   | |||||||  ||||   | ||||||||  |    |||||
Db       2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC 2412

Qy       2578 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2637
              ||  |||| |||||  ||  | |||       | |  ||  ||||||  |    |||||     ||
Db       2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472

Qy       2638 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
              |||  |   |||||  ||   ||        |       |   | ||  |  ||  ||||||
Db       2473 TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2532
```

```
RESULT 8
AY411749
LOCUS       AY411749                2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411749
VERSION     AY411749.1  GI:39767717
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2532
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1. .>2532
                     /locus_tag="HCM4327"
ORIGIN
```

Query Match 28.9%; Score 780.4; DB 9; Length 2532;
Best Local Similarity 60.1%; Pred. No. 5.1e-156;
Matches 1515; Conservative 0; Mismatches 876; Indels 129; Gaps 8;

```
Qy        307 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 366
              |||||  |  ||  | || ||| ||||||| || ||| |  | |||||||| ||||| |||
Db         13 GAGGTGCAGATCGAGGTGTCACGGCAGCAAGTGGAGGAACTCTTCGGGCTCGAGGACTAC 72

Qy        367 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 426
              |||||||||||||||||| ||||||||  |||||  || ||||||||||     |||||||||
Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGTCGCCGAGCCTACATC 132

Qy        427 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 486
              |||||  |||   |||||||||||||| || ||||||||  ||||||||||||||   || ||
Db        133 CGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCCAAGGAGGTACCCTTG 192

Qy        487 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 546
              ||  ||  |   ||  |  ||||  |||||  ||||||||||  |  |||   ||  ||||||||
Db        193 GATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAGTGCCTGTGGCTGAGGTGGAA 252

Qy        547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 606
              ||||||  |||  ||  ||  |  | |||||  | | ||||  | ||  |     ||||
Db        253 TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT 312

Qy        607 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 666
              ||  ||||  |||  |  | ||  |||||  |||||  | ||||||||||||||||||  |||
Db        313 GACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCCAACTACACCTGTGTG 372

Qy        667 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 726
              ||||||||  ||||||||     |||||  |||  ||  |  ||  |  ||||||||||  |||||
Db        373 GCCAAGAATATCGTGGCCAAGCGCCGGAGCACCACGGCCACAGTCATCGTCTATGTGAAT 432

Qy        727 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 786
              ||  ||  |||||  |  |||  |  ||||||||    |||    |    |||||||  ||  ||||||
Db        433 GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGCTGTGGCCGAGGCTGG 492

Qy        787 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 846
              |||||  ||  |  ||||  |||||||||  ||  ||  ||  |||||  ||  |||||  |||||||||
Db        493 CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTTCTGTGAG 552

Qy        847 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC 906
              ||  |||    |||||||  |||||  |||  |||||  ||||||||||  ||  ||    ||||  |
Db        553 GGACAGGCCTTCCAGAAGACAGCTTGCACCACCGTGTGCCCAGTGGATGGAGCGTGGACC 612

Qy        907 CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC 966
                |||||||||||||||  ||||||  |    ||  ||  |  ||||||||  |||||  ||||||
Db        613 GAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAGTGC 672

Qy        967 TCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC 1026
                |  ||  ||||||  ||||||||    || |||   ||  ||    ||  |||  ||
Db        673 ATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAG 732

Qy       1027 AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC---------- 1074
              |||||  ||   ||  ||  ||  ||  |  |  |         |   | |||
Db        733 AACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGCCAC 792

Qy       1075 ----------------------GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGC 1113
```



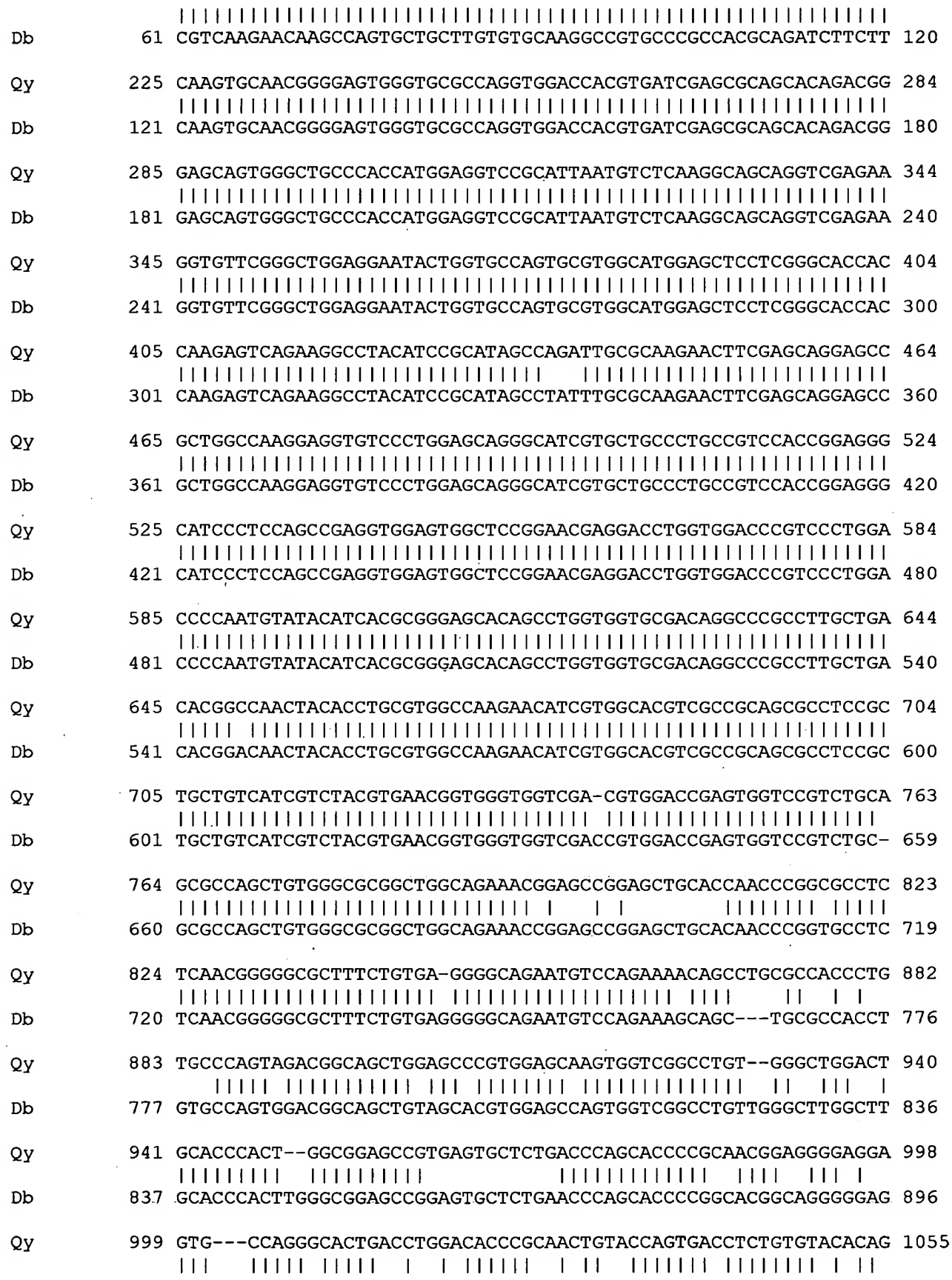


RESULT 9
BI758231
LOCUS BI758231 1034 bp mRNA linear EST 25-SEP-2001
DEFINITION 603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5', mRNA sequence.
ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1034)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11501 row: g column: 20
High quality sequence stop: 793.
FEATURES Location/Qualifiers
source 1. .1034
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200171"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."
ORIGIN

Query Match 28.5%; Score 768.6; DB 4; Length 1034;
Best Local Similarity 91.2%; Pred. No. 1.5e-153;
Matches 918; Conservative 0; Mismatches 74; Indels 15; Gaps 9;

Qy 105 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 164
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 60

Qy 165 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT 224



```
Db        897 GTGTGCCCAGGTCACTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCCAC 954

Qy       1056 TGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCG 1102
                | |||||||||||||||| |||||  |||||||||||||||| | |
Db        955 GGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATCCCGG 1001
```

RESULT 10
BC033727

```
LOCUS       BC033727                1532 bp    mRNA    linear   HTC 25-MAR-2004
DEFINITION  Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone
            IMAGE:5166762), containing frame-shift errors.
ACCESSION   BC033727
VERSION     BC033727.1  GI:21707230
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1532)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1532)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
```

Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 68 Row: i Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomeScan gene prediction
This clone has the following problem: frame shifted.

FEATURES Location/Qualifiers
source 1. .1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166762"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 27.3%; Score 736.6; DB 3; Length 1532;
Best Local Similarity 84.2%; Pred. No. 1.1e-146;
Matches 917; Conservative 0; Mismatches 4; Indels 168; Gaps 1;

```
Qy        292 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 671

Qy        352 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 731

Qy        412 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 471
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        732 CAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 791

Qy        472 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 851

Qy        532 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 911

Qy        592 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 971
```

```
Qy        652 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC 1031

Qy        712 ATCGTCTACGTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC 771
              ||||||||||||
Db       1032 ATCGTCTACGTG---------------------------------------------- 1043

Qy        772 TGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG 831

Db       1044 ---------------------------------------------------------- 1043

Qy        832 GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA 891

Db       1044 ---------------------------------------------------------- 1043

Qy        892 GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1044 GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG 1103

Qy        952 CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1104 CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT 1163

Qy       1012 GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAG 1071
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1164 GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAG 1223

Qy       1072 GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTT 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1224 GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTT 1283

Qy       1132 GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1284 GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG 1343

Qy       1192 TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC 1251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1344 TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC 1403

Qy       1252 CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTC 1311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTC 1463

Qy       1312 TGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 TGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC 1523

Qy       1372 CCCCTGGGT 1380
              |||||||||
Db       1524 CCCCTGGGT 1532
```

RESULT 11
AI951556
LOCUS AI951556 788 bp mRNA linear EST 09-MAR-2000

DEFINITION wv36f04.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:2531647 3' similar to TR:O08721 O08721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA sequence.
ACCESSION AI951556
VERSION AI951556.1 GI:5743866
KEYWORDS EST.
SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 788)
  AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1125 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 446.
FEATURES Location/Qualifiers
  source 1. .788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2531647"
/tissue_type="fibrotheoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Ov18"
/note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGCGACATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Query Match 27.3%; Score 735.2; DB 1; Length 788;
Best Local Similarity 96.5%; Pred. No. 2e-146;
Matches 749; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1424 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1483
Db 1 CCGAGGAGTTCGTCTCCCGCCTCTCCGCCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60

Qy 1484 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1543

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 120

Qy        1544 CAGGTATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 1603
                ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CAGGAATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT 180

Qy        1604 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 1663
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGC 240

Qy        1664 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG 1723
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         241 TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTTCTGCTCACCCGGCCAGTCATCCTGG 300

Qy        1724 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 1783
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT 360

Qy        1784 CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 1843
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT 420

Qy        1844 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG 1903
                |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||  ||
Db         421 ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATG 480

Qy        1904 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 1963
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG 540

Qy        1964 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCC 2023
                |||||||||||||||||||||||||||||||| | |||||||||||||||||||| |
Db         541 CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGCATGACACTC 600

Qy        2024 ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG 2083
                |||||||||||||  |  |||||||||||||||||||||||  ||||||||||||||||
Db         601 ACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGG 660

Qy        2084 AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATG 2143
                ||||||  ||| ||||||||||||||||||||||||||||||||||||  ||||||||||
Db         661 AGCCACTGGTACTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATNCATCCACGATG 720

Qy        2144 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTT 2199
                |||||||||  ||    | ||||||||| ||||||||||||  ||||||| || |
Db         721 TGCCCAGCTNCCNTGTGGAGAGTAAGCTTCTTGTCAGCTACCCAGAGATCCNCTAT 776
```

```
RESULT 12
BX348193/c
LOCUS       BX348193                796 bp    mRNA    linear   EST 08-APR-2004
DEFINITION  BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
            cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION   BX348193
VERSION     BX348193.2  GI:46286231
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 5, 2003 this sequence version replaced gi:30367258.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3239.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0BAF004ZD01_AF00293_1&c=3239.r
            .
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DB008YE02"
                     /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
                     /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             26.8%;  Score 721.6;  DB 5;  Length 796;
  Best Local Similarity   96.4%;  Pred. No. 1.7e-143;
  Matches  758;  Conservative    0;  Mismatches   26;  Indels    2;  Gaps    2;

Qy       1634 GGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTG 1693
              | |  |||  | | ||||||||||||||| |||||||||| |||||||||| | ||||||
Db        795 GTTGCCCCAAACCGGGCTGTCAGACCCTGTTGAGTCCCATGGTTAGCTGTGAA-CCCCTG 737

Qy       1694 GCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTG 1753
              ||||||||||||| | |||||||||||| |||||||||||||||||||||||||||||
Db        736 GCGTCCTGCTCACCGGCCCAGTCATCCTGGGTATGGACCACTGTGGGGAGCCCAGCCCTG 677

Qy       1754 ACAGCTGGAGCCTGCGCCTC-AAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTG 1812
              ||||||| |||||||||| ||||| |||||||||||| |||||||||| |||| |||
Db        676 ACAGCTGAGGCCTGCGCCTCAAAAAAACAGTCGTGCGAGGACAGCTGGGAGTATGTCCTG 617

Qy       1813 CACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 1872
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        616 CACCTGGGCGAGNAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGC 557

Qy       1873 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        556 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 497

Qy       1933 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTAC 437

Qy       1993 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 377

Qy       2053 GAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT 2112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT 317

Qy       2113 TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC 2172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC 257

Qy       2173 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 2232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 197

Qy       2233 CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 2292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 137

Qy       2293 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG 77

Qy       2353 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTG 2412
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||   |
Db         76 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGNCCCTGGT 17

Qy       2413 GGCCCC 2418
              || |||
Db         16 GGGCCC 11
```

```
RESULT 13
BI818609
LOCUS       BI818609                 818 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603033362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
            mRNA sequence.
ACCESSION   BI818609
VERSION     BI818609.1  GI:15929902
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11434 row: 1 column: 16
High quality sequence stop: 744.

FEATURES Location/Qualifiers
source 1. .818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174559"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.5%; Score 713.4; DB 4; Length 818;
Best Local Similarity 97.6%; Pred. No. 9.4e-142;
Matches 745; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363
```

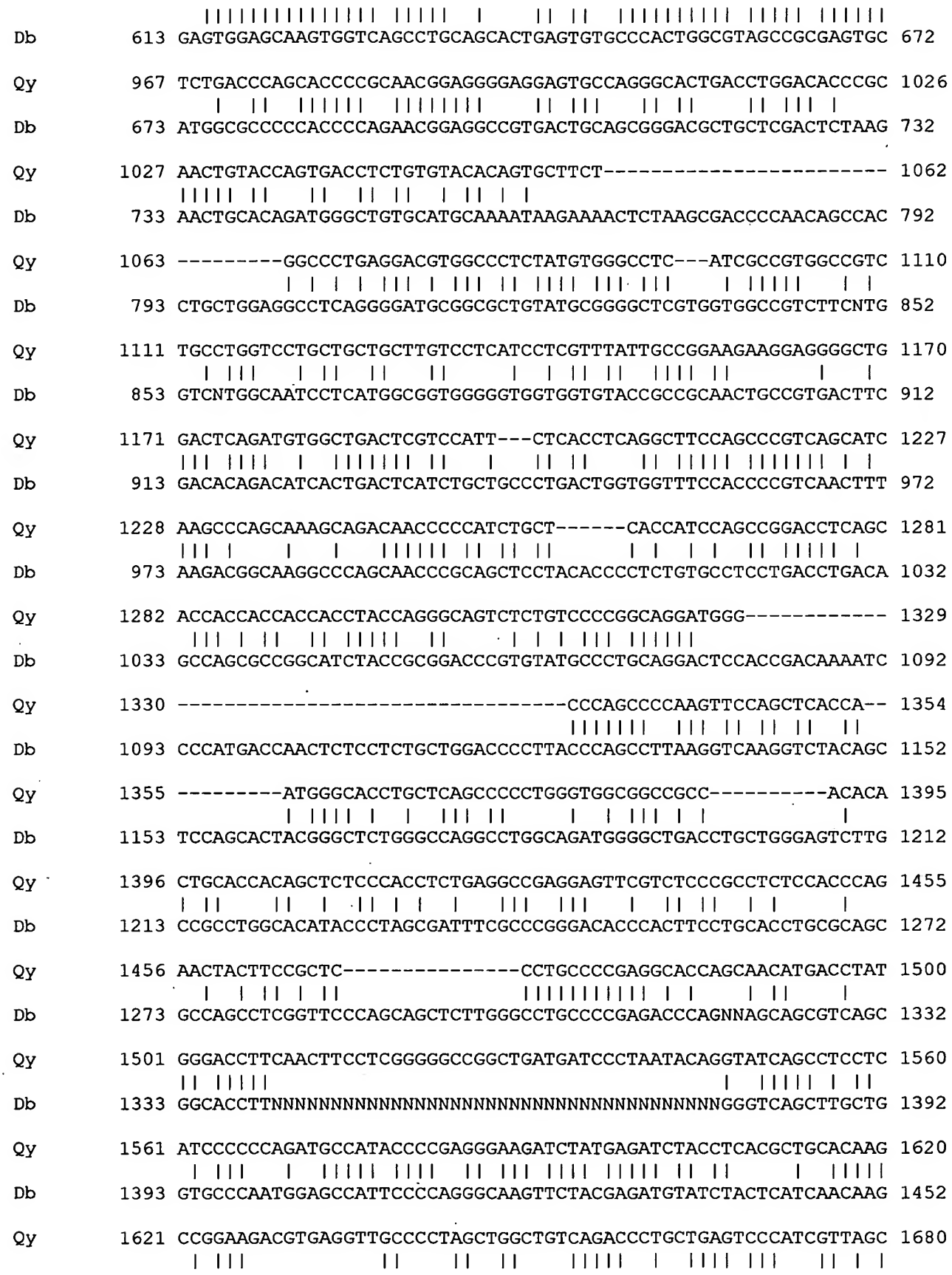
```
Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 603

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 663

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTC-GCCGCAGCGCCTCCGCTGCTGTCATCGTCTA 719
              |||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||
Db        664 TGCGTGGCCAAGAACATCGTGGCACGTCAGCCGCAGGGCCTCCGCTGCTGTCATCGTCTA 723

Qy        720 CGTGAACGGTGGGTGGTCGACG-TGGACCGAGTGGTCCGTCTG 761
              |||| ||||  | |||    || |||| | ||||||| | |||
Db        724 CGTGGACGGCAGCTGGAGCCCGTTGGAGCCAGTGGTCGGGCTG 766
```

```
RESULT 14
AY411748
LOCUS       AY411748               2532 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY411748
VERSION     AY411748.1  GI:39767716
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 2532)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
```

```
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .2532
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>2532
                     /locus_tag="HCM4327"
ORIGIN
```

Query Match 25.1%; Score 678.2; DB 9; Length 2532;
Best Local Similarity 54.4%; Pred. No. 3.7e-134;
Matches 1370; Conservative 0; Mismatches 1021; Indels 129; Gaps 9;

```
Qy        307 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 366
              |||||  |  ||  | || ||  |||||||||| ||| || | || |||||||||||  |||
Db         13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72

Qy        367 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC 426
              |||||||||||||||| ||||||||  ||||||||||||||||||      |||||| ||
Db         73 TGGTGCCAGTGCGTGGCCTGGAGCTCTGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC 132

Qy        427 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 486
              |||||  |||
Db        133 CGCATCGCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 192

Qy        487 GAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 546
                                                                  |||||
Db        193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGGAA 252

Qy        547 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 606
              ||||||  ||| |||||  |  | |||||  ||| |||| ||||  |     ||||
Db        253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312

Qy        607 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 666
              || |||| |||  |  | || ||||||||||  | ||||| |||||||| |||||||||
Db        313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372

Qy        667 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 726
              ||||||||||||||||    ||||| ||| || | ||  | ||||||||||||||||
Db        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432

Qy        727 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 786
              || || ||||| |  ||| | ||||||||   ||||  |  | |||||||  |   ||||
Db        433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCGCCCTGCTCCAACCGCTGTGGNNGANNCTGG 492

Qy        787 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 846
              ||||| || | ||||| |||||||||||| || || |||||||| ||     |||| |||
Db        493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCCCCACTCAACGGAGGGNNNNTCTGCGAG 552

Qy        847 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC 906
              || ||                                        | || ||    ||||
Db        553 GGCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCGATGGGGCGTGGACG 612

Qy        907 CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC 966
```



```
Db       1453 GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCAGTGACC 1512

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
              |||||||||  | |||  |||||||    ||  ||  ||||||||   | |||   ||||||||
Db       1513 TGTGGACCCACAGGCCTCCTGCTGTGCCACCCCGTCATCCTCACCATGCCCCACTGTGCC 1572

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
              ||   |||  |       |||||  |  |  |  |||||  |  |||  |    | |||||  ||||
Db       1573 GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG 1632

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
              |||||  |||  ||  |||||  ||||||  |  |  .   |    | ||  ||||||||||||||||
Db       1633 GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACACCCTGCTACTGCCAGCTGGAG 1692

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
               ||||  |||||   ||  ||  |     ||  |||||||||   ||   |    |   |||  |||  ||
Db       1693 CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC 1752

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
                    |       | |  |       |||  ||||    | ||  ||  ||  |    |||||||
Db       1753 TATTCCCNNNNNNCAGTCNNNNNGCTCCNGCTGGCCGTCTTNGCCCCTGCCCTCTGCACC 1812

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
              |||||  ||       |  ||||||||||||||||                          |||
Db       1813 TCCCTGGANNNNNGCCTCCGGGTCTACTGCCTGGNNNNNNNNNNNNNNNNNNNNNNGAG 1872

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
              |||  ||  ||||||||   |    |||||  |||  |   ||  |   |||||||        ||
Db       1873 GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG 1932

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
              |||||||||||||||||||||||||||||  |||  ||||  ||   |  |||     |||   |||
Db       1933 TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG 1992

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
              |  |||  |||||  ||  |  ||  |||||||||||||||||  |  ||||||  ||||  |||||
Db       1993 AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTGTCACATTTGGAGTGGCAGC 2052

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
              |||   |   |  |  ||||||||  ||||||||||||  |    ||||      |  |  |   ||  ||
Db       2053 CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC 2112

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
               ||||||||  |  ||  |||  ||||  |||||  ||  ||  ||.  ||||   |||     |   |    |
Db       2113 ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC 2172

Qy       2341 AACATCACCAAG---GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2397
              |    |  |  ||        |  | |      ||     ||        |     ||
Db       2173 ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT 2232

Qy       2398 GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2457
              |||  |   |||  |  ||||  ||    ||||||||||||||   |    |||  ||  |||||||||
Db       2233 GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA 2292

Qy       2458 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2517
                    ||||||  ||   |  |||       |||||     ||||||||||    |  ||  |||||
Db       2293 TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG 2352
```

```
Qy       2518 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2577
              |||    ||||| |  | || |  | |||||||| ||||  | |||||||| |   |||||
Db       2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC 2412

Qy       2578 CTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCA 2637
              ||  |||| ||||| || | |||      | | ||  |||||| |   |||||    ||
Db       2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472

Qy       2638 GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
               |||  |   ||||| ||   ||        |       |   | || |  || ||||||
Db       2473 TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2532


RESULT 15
BF311804
LOCUS       BF311804                 934 bp    mRNA    linear   EST 21-NOV-2000
DEFINITION  601897316F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4126706 5',
            mRNA sequence.
ACCESSION   BF311804
VERSION     BF311804.1  GI:11259566
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 934)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM1016  row: p  column: 03
            High quality sequence stop: 707.
FEATURES             Location/Qualifiers
     source          1. .934
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4126706"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_19"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
```

ORIGIN

Query Match 24.9%; Score 672.6; DB 2; Length 934;
Best Local Similarity 93.3%; Pred. No. 5.1e-133;
Matches 747; Conservative 0; Mismatches 49; Indels 5; Gaps 4;

```
Qy       1071 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 1130
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 61

Qy       1131 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 1190
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 121

Qy       1191 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 1250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 181

Qy       1251 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 1310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 241

Qy       1311 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 1370
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 301

Qy       1371 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 361

Qy       1431 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 421

Qy       1491 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT 1550
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        422 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT 481

Qy       1551 CAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 1610
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CAG-CTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 540

Qy       1611 GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 1670
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        541 GCTGCACAAGCCGGAAGACGTGAGGTTG-CCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 599

Qy       1671 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCC--TGGCTATG 1728
             |||||||||||||||||||||||||||||||| |||| ||||||||| |   |||
Db        600 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCAACCGGGCAGTCATCCCTGGCTAATG 659

Qy       1729 GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC 1788
             ||||      |||    |  |||||||| |||||||| ||   |||| |     ||||
Db        660 GACCCACTGTGGGGAAGCCAGCCTGACAGTTGGAGCCTGGGCTCAAAAAGGAGTCGGTGC 719

Qy       1789 GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC 1848
             |||||||||||||| ||||||||| ||  |||  || ||| | |||||  |||||| ||
Db        720 GAGGGCAGCTGGGA-GATGTGCTGGACTGGGGGAAGAAGGGGGCCTCCACCCTCTATTAT 778
```

```
Qy       1849 TGCCAGCTGGAGGCCAGTGCC 1869
                || |||  |||||| ||  ||
Db        779 TGGCAGTGGGAGGCAAGGCCC 799
```

Search completed: March 6, 2005, 10:10:35
Job time : 8489.11 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:11:26 ; Search time 11456.2 Seconds
(without alignments)
11407.261 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggcccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2697	100.0	2752	6	AX449572	AX449572 Sequence
2	2687.4	99.6	2697	6	AX451652	AX451652 Sequence
3	2621.4	97.2	2881	6	AX527916	AX527916 Sequence

4	2435.4	90.3	2784	6	CQ730306	CQ730306 Sequence
5	2343	86.9	3580	6	AX367094	AX367094 Sequence
6	2297	85.2	3992	10	MMU487852	AJ487852 Mus muscu
7	2252.2	83.5	2697	6	AX268596	AX268596 Sequence
8	2252.2	83.5	2697	10	RNU87305	U87305 Rattus norv
9	2252.2	83.5	3014	6	BD057524	BD057524 Netrin re
10	1957.4	72.6	3844	10	BC058084	BC058084 Mus muscu
11	1625.4	60.3	2688	9	BC009333	BC009333 Homo sapi
12	1552.4	57.6	1787	6	BD057525	BD057525 Netrin re
13	1302.8	48.3	9700	6	AX054976	AX054976 Sequence
14	992	36.8	9299	10	MMU72634	U72634 Mus musculu
15	988.8	36.7	2962	5	AY187310	AY187310 Gallus ga
16	986	36.6	9328	10	AB118026	AB118026 Rattus no
17	970	36.0	3646	6	CQ881060	CQ881060 Sequence
18	970	36.0	3646	9	AF055634	AF055634 Homo sapi
19	942.4	34.9	3770	9	AY126437	AY126437 Homo sapi
20	936.2	34.7	2860	6	AX686445	AX686445 Sequence
21	936.2	34.7	2860	6	AX686447	AX686447 Sequence
22	912.8	33.8	3672	10	BC048162	BC048162 Mus muscu
23	912.8	33.8	3672	10	BC057560	BC057560 Mus muscu
24	912.2	33.8	2895	6	AX512281	AX512281 Sequence
25	904.2	33.5	2986	6	CQ881052	CQ881052 Sequence
26	904.2	33.5	3933	6	CQ850929	CQ850929 Sequence
27	904.2	33.5	3933	9	AK128132	AK128132 Homo sapi
28	904.2	33.5	3935	9	AB096256	AB096256 Homo sapi
29	901	33.4	3884	6	AR528525	AR528525 Sequence
30	901	33.4	3884	6	AX464012	AX464012 Sequence
31	901	33.4	3884	9	AY358351	AY358351 Homo sapi
32	892.4	33.1	2995	6	AX497288	AX497288 Sequence
33	892	33.1	4294	10	AK122575	AK122575 Mus muscu
34	874.6	32.4	3788	10	MMU487853	AJ487853 Mus muscu
35	852.4	31.6	2838	10	RNU87306	U87306 Rattus norv
36	843.6	31.3	2625	6	CQ721377	CQ721377 Sequence
37	841.4	31.2	2831	6	BD057526	BD057526 Netrin re
38	764.8	28.4	813	6	AX054892	AX054892 Sequence
39	751.4	27.9	4330	5	AY744919	AY744919 Petromyzo
40	738.4	27.4	2230	6	CQ845766	CQ845766 Sequence
41	738.4	27.4	2230	9	AK131380	AK131380 Homo sapi
42	717.2	26.6	2832	5	AY099459	AY099459 Xenopus l
43	668.4	24.8	2612	6	CQ881064	CQ881064 Sequence
44	623.2	23.1	2661	6	AX800717	AX800717 Sequence
45	623.2	23.1	2868	6	AX800719	AX800719 Sequence

ALIGNMENTS

RESULT 1
AX449572
LOCUS AX449572 2752 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0210216.
ACCESSION AX449572
VERSION AX449572.1 GI:21698195
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Padigaru,M., Mezes,P., Mishra,V., Burgess,C., Casman,S.,
            Grosse,W.M., Alsobrook,J.P., Lepley,D.M., Gerlach,V.L.,
            Macdougall,J.R. and Smithson,G.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0210216-A 1 07-FEB-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .2752
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2697;  DB 6;  Length 2752;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2697;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 945

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 1005

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1065

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1125

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1185

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1245

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1305

Qy       1261 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1365

Qy       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1425

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1485
```

```
Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1545

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1605

Qy       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1665

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1725

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1785

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1845

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1905

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1965

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2025

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2085

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2145

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2325
```

```
Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
```

```
RESULT 2
AX451652
LOCUS       AX451652                2697 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 1 from Patent WO0233080.
ACCESSION   AX451652
VERSION     AX451652.1  GI:21698587
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Koehler,R.H.
  TITLE     Regulation of human netrin binding membrane receptor unc5h-1
  JOURNAL   Patent: WO 0233080-A 1 25-APR-2002;
            Bayer Aktiengesellschaft (DE)
FEATURES             Location/Qualifiers
     source          1. .2697
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.6%;  Score 2687.4;  DB 6;  Length 2697;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2691;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
```

```
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC 900

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC 1080

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260

Qy       1261 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1320

Qy       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440

Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC 1560

Qy       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
```

```
Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
```

```
Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697


RESULT 3
AX527916
LOCUS       AX527916                2881 bp    DNA     linear   PAT 21-NOV-2002
DEFINITION  Sequence 1 from Patent WO0229038.
ACCESSION   AX527916
VERSION     AX527916.1  GI:25172359
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Herrmann,J.L., Rastelli,L. and Shimkets,R.A.
  TITLE     Novel proteins and nucleic acids encoding same and antibodies
            directed against these proteins
  JOURNAL   Patent: WO 0229038-A 1 11-APR-2002;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .2881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             97.2%;  Score 2621.4;  DB 6;  Length 2881;
  Best Local Similarity   98.9%;  Pred. No. 0;
  Matches 2673;  Conservative    0;  Mismatches   21;  Indels    9;  Gaps    3;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
```

```
Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              |||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
                ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
                ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
              |||
Db       2784 TGA 2786
```

RESULT 4
CQ730306
LOCUS CQ730306 2784 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16240 from Patent WO02068579.
ACCESSION CQ730306
VERSION CQ730306.1 GI:42303801

KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 16240 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source 1. .2784
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 90.3%; Score 2435.4; DB 6; Length 2784;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 6; Indels 177; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

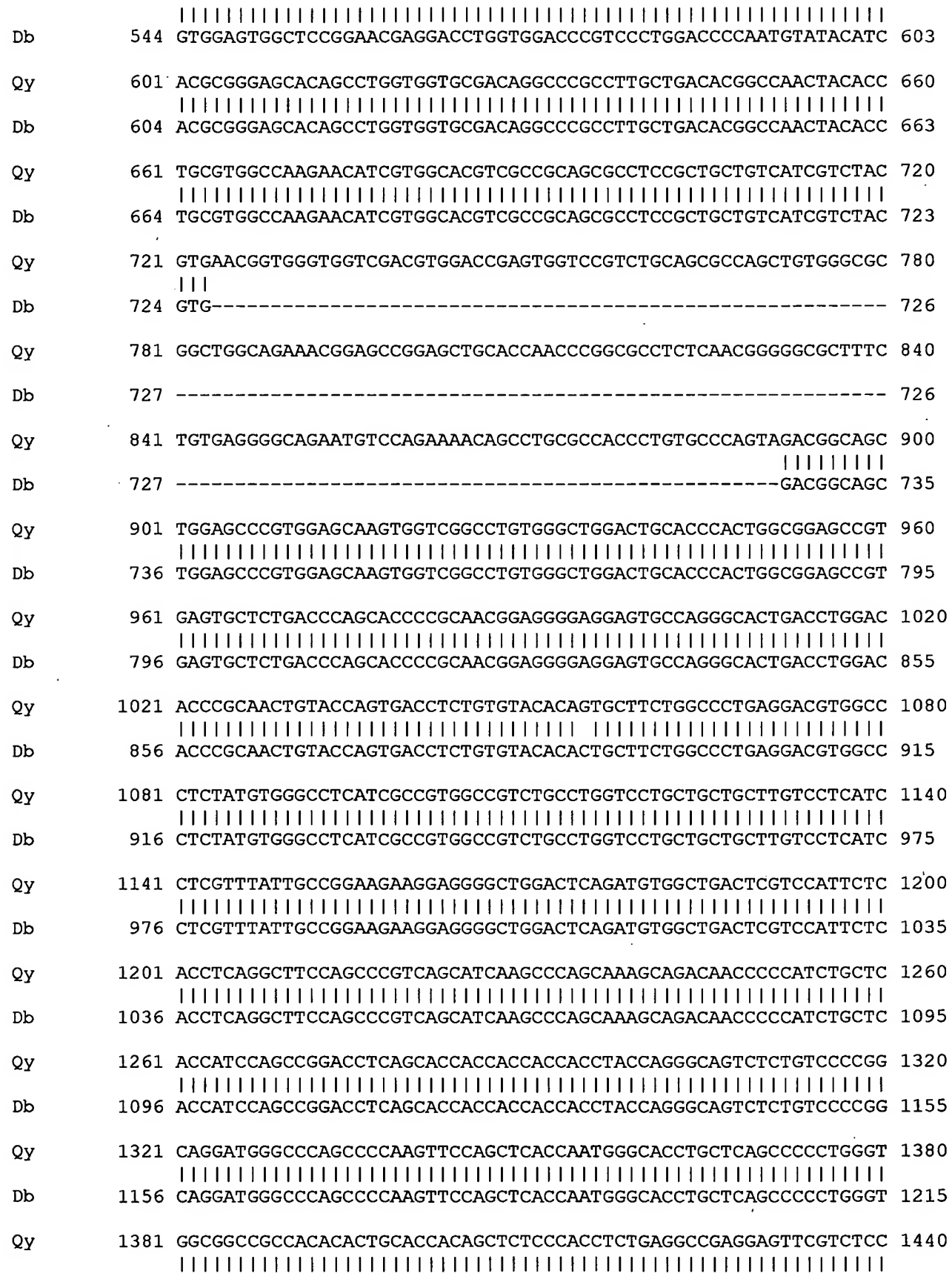
```
Db        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720

Qy        721 ------------------------------------------------------------ 720

Db        721 GGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGCCCCTGGGCAGTGAC 780

Qy        721 ---------------------------GTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 753
                                          ||||||||||||||||||||||||||||||||
Db        781 ATGTGGCTGTCCTTCTCTGTCCGGCCAGTGAACGGTGGGTGGTCGACGTGGACCGAGTGG 840

Qy        754 TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 813
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC 900

Qy        814 CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC 873
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC 960

Qy        874 GCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG 933
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCACCCTGTGCCCAGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG 1020

Qy        934 CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG 993
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG 1080

Qy        994 GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC 1140

Qy       1054 A-------------------------------------------GTGCTTCTGGCCCT 1068
              |                                            ||||||||||||
Db       1141 AACTCCTACACCCCTGCCCCCACCAAGGCCATGCTGTCTCCCGCAGCTGCTTCTGGCCCT 1200

Qy       1069 GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTG 1260

Qy       1129 CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC 1320

Qy       1189 TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC 1380
```

```
Qy       1249 CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGT 1308
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGT 1440

Qy       1309 CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC 1368
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC 1500

Qy       1369 AGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG 1428
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG 1560

Qy       1429 GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC 1620

Qy       1489 AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGT 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGA 1680

Qy       1549 ATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ATCAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC 1740

Qy       1609 ACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGT 1668
             ||||||||||||||||||||||
Db       1741 ACGCTGCACAAGCCGGAAGACGTG---------------------------------- 1764

Qy       1669 CCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG 1728
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1765 ---------AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG 1815

Qy       1729 GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1816 GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC 1875

Qy       1789 GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1876 GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC 1935

Qy       1849 TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1936 TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG 1995

Qy       1909 GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1996 GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG 2055

Qy       1969 GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGAT 2028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2056 GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGAT 2115

Qy       2029 GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCA 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2116 GCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCA 2175
```

```
Qy       2089 CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCC 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2176 CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCC 2235

Qy       2149 AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2236 AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC 2295

Qy       2209 TGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC 2268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2296 TGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC 2355

Qy       2269 ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC 2328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2356 ACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC 2415

Qy       2329 AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2416 AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT 2475

Qy       2389 GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2476 GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG 2535

Qy       2449 CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2536 CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG 2595

Qy       2509 GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA 2568
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2596 GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA 2655

Qy       2569 GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTG 2628
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2656 GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTG 2715

Qy       2629 GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT 2688
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2716 GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT 2775

Qy       2689 GAGTGCTGA 2697
             |||||||||
Db       2776 GAGTGCTGA 2784
```

RESULT 5
AX367094
LOCUS AX367094 3580 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 13 from Patent WO0198354.
ACCESSION AX367094
VERSION AX367094.1 GI:18855296
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1

```
  AUTHORS   Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
            Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
            Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
            Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
            Sanjanwala,M.S.
  TITLE     Receptors
  JOURNAL   Patent: WO 0198354-A 13 27-DEC-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3580
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 6052371CB1"
ORIGIN

  Query Match             86.9%;  Score 2343;  DB 6;  Length 3580;
  Best Local Similarity   93.6%;  Pred. No. 0;
  Matches 2524;  Conservative    0;  Mismatches    5;  Indels  168;  Gaps    1;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 63

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 123

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 183

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 243

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 303

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 363

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 423

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        424 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 483

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 543

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
```



```
Db      1216 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1275

Qy      1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1276 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1335

Qy      1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      1336 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC 1395

Qy      1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1396 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1455

Qy      1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1456 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1515

Qy      1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1516 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1575

Qy      1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1576 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1635

Qy      1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1636 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1695

Qy      1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1696 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1755

Qy      1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1756 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1815

Qy      1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1816 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 1875

Qy      2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1876 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 1935

Qy      2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1936 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 1995

Qy      2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1996 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2055

Qy      2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2056 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2115
```

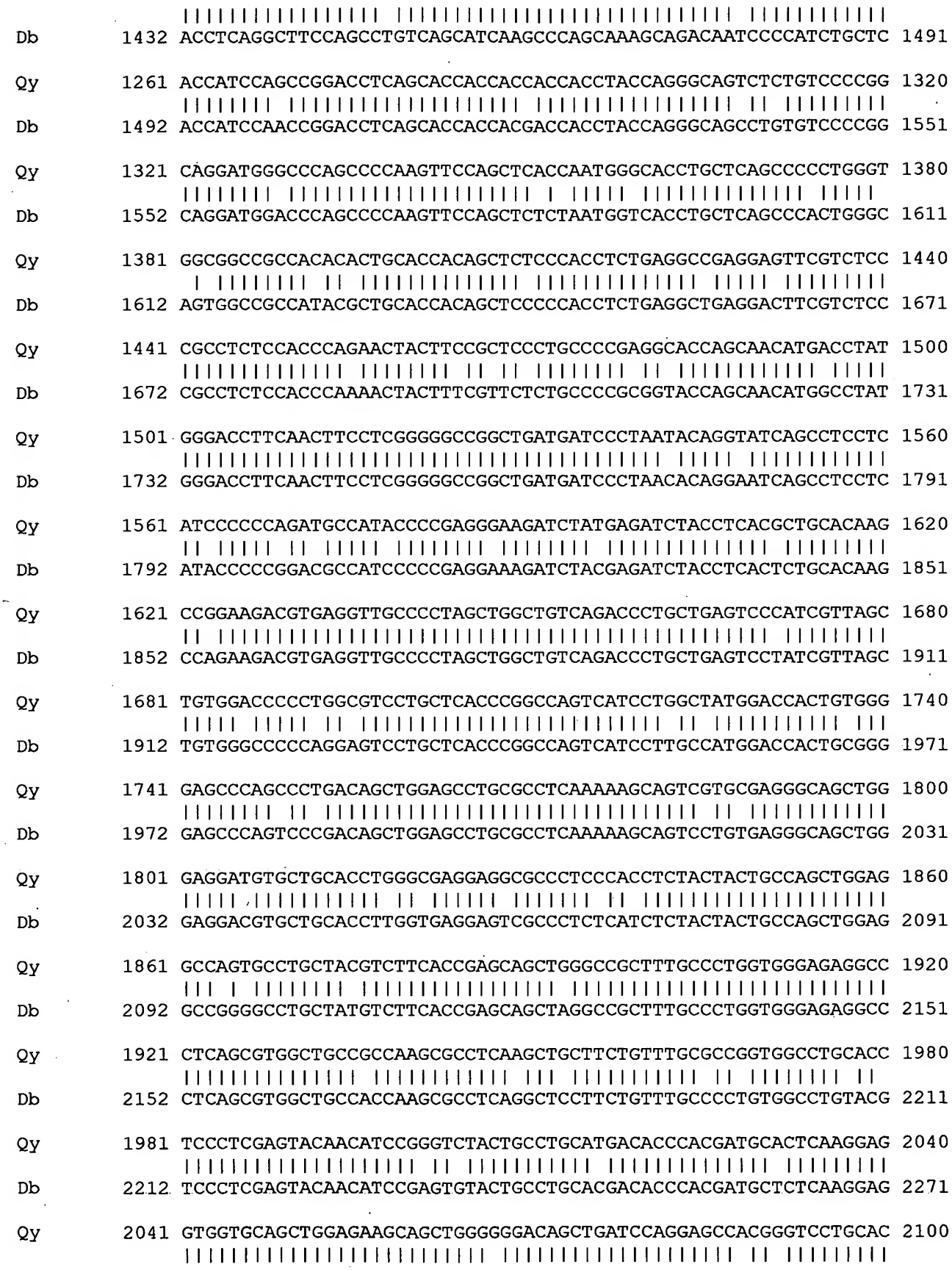
```
Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2116 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2175

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2176 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2235

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2236 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2295

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2296 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2355

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2356 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2415

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2416 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2475

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2476 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2532
```

```
RESULT 6
MMU487852
LOCUS       MMU487852               3992 bp    mRNA    linear   ROD 24-SEP-2002
DEFINITION  Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).
ACCESSION   AJ487852
VERSION     AJ487852.1  GI:22035783
KEYWORDS    netrin receptor Unc5h1; Unc5h1 gene.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Engelkamp,D.
  TITLE     Cloning of three mouse Unc5 genes and their expression patterns at
            mid-gestation
  JOURNAL   Mech. Dev. 118 (1-2), 191-197 (2002)
  MEDLINE   22239710
   PUBMED   12351186
REFERENCE   2  (bases 1 to 3992)
  AUTHORS   Engelkamp,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for
            Brain Research, Deutschordenstrasse 46, Frankfurt 60528, GERMANY
FEATURES             Location/Qualifiers
     source          1. .3992
                     /organism="Mus musculus"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:10090"
     gene            1. .3992
                     /gene="Unc5h1"
     CDS             232. .2928
                     /gene="Unc5h1"
                     /codon_start=1
                     /product="netrin receptor Unc5h1"
                     /protein_id="CAD32250.1"
                     /db_xref="GI:22035784"
                     /db_xref="GOA:Q8K1S4"
                     /db_xref="UniProt/TrEMBL:Q8K1S4"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALESEG
                     GVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
                     AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN
```

```
  Query Match             85.2%;  Score 2297;  DB 10;  Length 3992;
  Best Local Similarity   90.7%;  Pred. No. 0;
  Matches 2447;  Conservative    0;  Mismatches  250;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| | || |||||
Db        232 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTT 291

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        292 CGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCA 351

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| ||||||||||||||
Db        352 GACCTGCTGCCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCC 411

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              |||||||| |||||||||| |||||||||| ||||||||||||||||||||||||||||
Db        412 GTGCTGCTGGTGTGCAAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAA 471

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||| |||||||| || ||||| ||||| |||||||| ||||| ||||| ||  ||||
Db        472 TGGGTTCGCCAGGTCGATCACGTCATCGAACGCAGCACTGACGGCAGCAGCGGATTGCCA 531

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              |||||||||||||| || || || |||||||||||||||||||| ||||| |||||||||
Db        532 ACCATGGAGGTCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAG 591
```

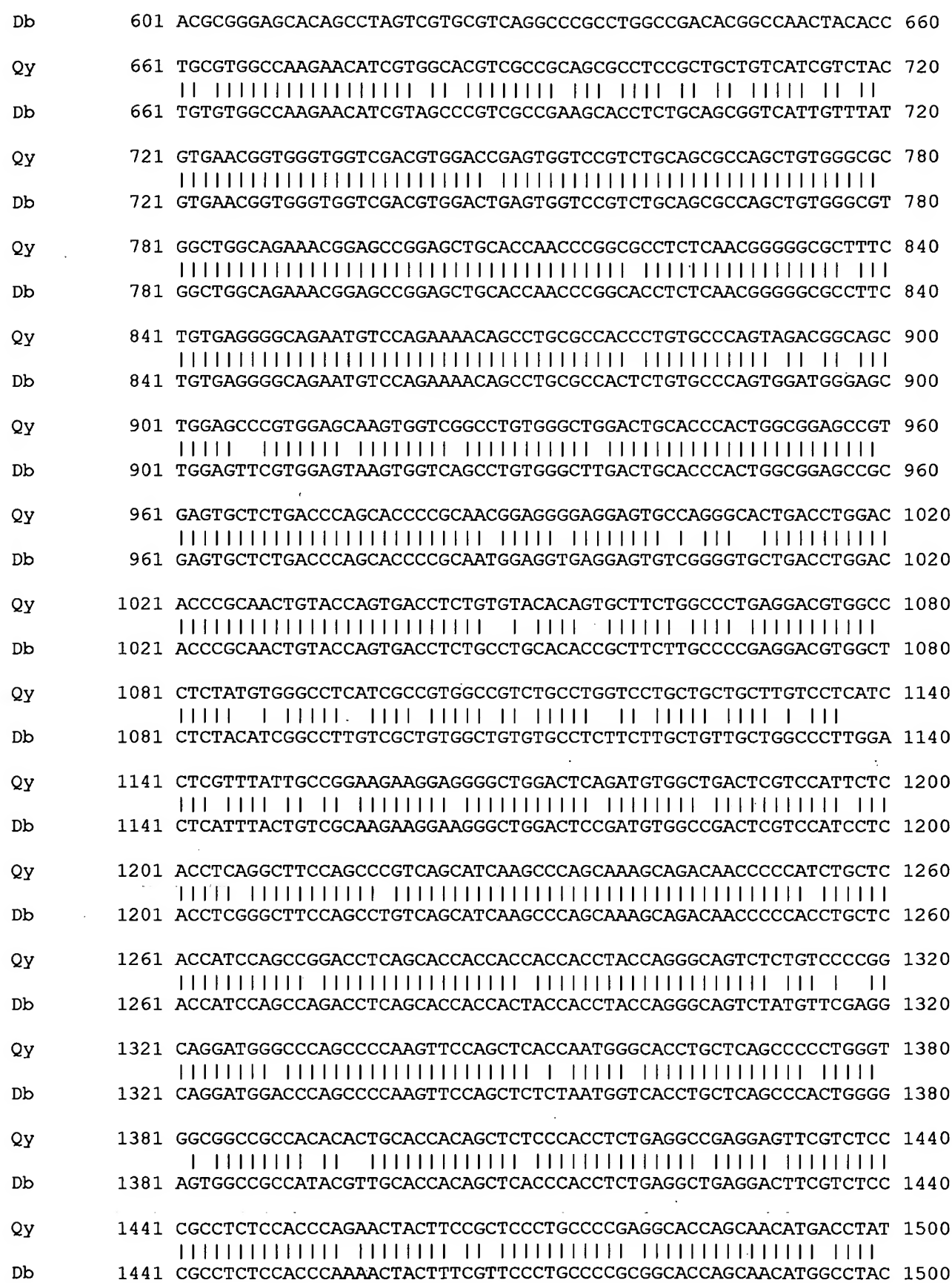
```
Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              || |||||||||||||| |||||||||||||||| || |||||||| || |||||||||
Db        592 GAGTACTGGTGCCAGTGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCC 651

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||| || |||   |||||||||||||| |||||||||||||||||||||||| |||
Db        652 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTG 711

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              || |||||||| ||||| ||||| || || || || ||||| || ||||| ||||| |||
Db        712 TCACTGGAGCAAGGCATTGTGCTACCTTGTCGCCCCCCGGAAGGAATCCCCCCAGCTGAG 771

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              |||||||||||||| || |||||||| |||||||| ||||| |||||||||| ||||||
Db        772 GTGGAGTGGCTCCGAAATGAGGACCTCGTGGACCCCTCCCTCGACCCCAATGTGTACATC 831

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              || |||||||||||||| || ||||| |||||||||| || ||||| ||||||||||||
Db        832 ACACGGGAGCACAGCCTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTGCCAACTACACC 891

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              |||||||||||||||||||||||| |||||||| |||||||| || || ||||| || ||
Db        892 TGCGTGGCCAAGAACATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTAT 951

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        952 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGGCGT 1011

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db       1012 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 1071

Qy        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 900
              |||||||||||||||||||||||||||||||||||||| |||||||||| || || |||
Db       1072 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC 1131

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              |||||||| ||||| |||||||| ||||| ||||| ||||||||||||||||||||||
Db       1132 TGGAGCCCATGGAGTAAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGG 1191

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              |||||||| |||||||| |||||||||||||| |||||||||| |||  ||||||||||
Db       1192 GAGTGCTCCGACCCAGCGCCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGAC 1251

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              |||||||||||||||||||||||||  | ||||   |||| ||||| ||||||||||
Db       1252 ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCT 1311

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              |||||  | |||||| |||||||||||| ||||| || |||||||||| |||||| ||
Db       1312 CTCTACATCGGCCTCGTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTCGTC 1371

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||| | || ||||| |||||||| || |||||||||| |||||||||| ||||| ||
Db       1372 CTCATCTACTGCCGCAAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCATCCTT 1431

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
```

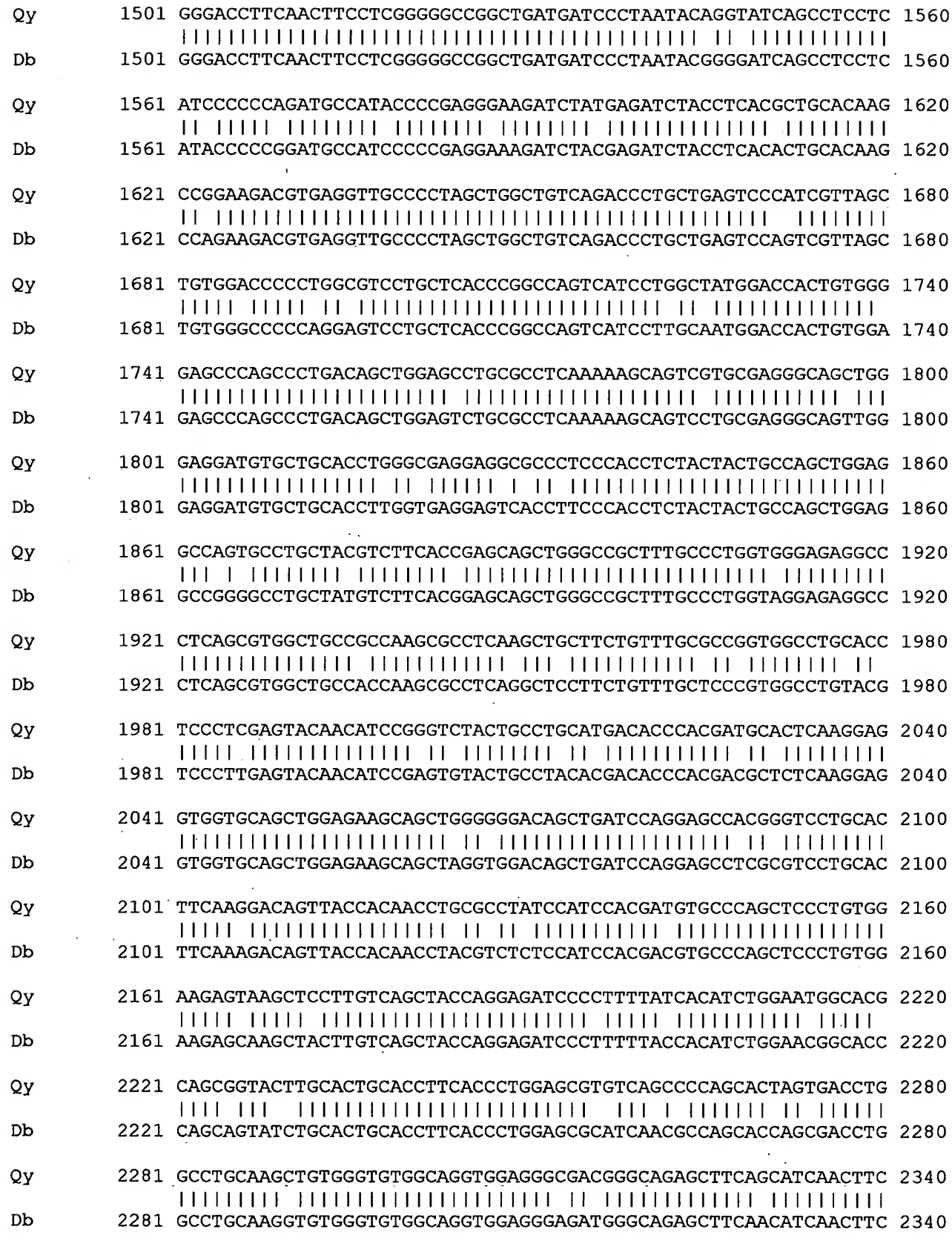


```
Db       2272 GTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCAC 2331

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
              ||||| |||||||||||||||| || || |||||||||| ||||||||||||||||||
Db       2332 TTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGG 2391

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
              ||||| |||||||||||||||||||||||||| ||||| ||||||||||||||||
Db       2392 AAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACT 2451

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
              |||| |||  ||||||||||||||||||||||| ||||   ||||||| || ||||||
Db       2452 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTG 2511

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
              |||||||| |||||||||||||||||||| || || |||||||||| ||||||||
Db       2512 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTT 2571

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
              |||||||| |||||||| |||||||||| |||||||||||||||||||| |||||||
Db       2572 AACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTC 2631

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
Db       2632 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATT 2691

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
               |||||||||||||||||| || ||||| |||||||||| ||||| ||||||||||
Db       2692 ACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTT 2751

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||
Db       2752 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2811

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
              ||||| |||||||||||||||||||||||||||||||| ||||||||| ||||| |||
Db       2812 AACCTATGGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTG 2871

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
              || |||||||||||||||| |||||||||||||| ||||| ||||| |||||||||
Db       2872 GCCGGACTGGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGA 2928
```

```
RESULT 7
AX268596
LOCUS       AX268596               2697 bp    DNA     linear   PAT 29-OCT-2001
DEFINITION  Sequence 15 from Patent WO0175440.
ACCESSION   AX268596
VERSION     AX268596.1  GI:16541710
KEYWORDS    .
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and
```

```
            Pratt,J.Y.
  TITLE     Schizophrenia related genes
  JOURNAL   Patent: WO 0175440-A 15 11-OCT-2001;
            WELFIDE CORPORATION (JP)
FEATURES             Location/Qualifiers
     source          1. .2697
                     /organism="Rattus sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10118"
ORIGIN

  Query Match             83.5%;  Score 2252.2;  DB 6;  Length 2697;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches 2419;  Conservative    0;  Mismatches  278;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||||| |||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || |||||||||||||||||||||| |||||||| |||||||| ||||||||||
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| ||||||||||||||
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||| || | |||||||||| ||||| ||||| ||||||||||||||||||| |||||
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||| |||||||| || ||||| || || |||||||| ||| | ||||| ||  ||||
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              |||||||||||||| || || || || |||||||||| ||||| ||||| |||||||||
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              |||||||||||||||| |||||||||||||||||| |||||||| ||||||||||||
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||| || |||   |||||||||||||| |||||||||| |||||||||| |||
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              || |||||||| ||||| || || || || || || || || || ||||| ||||| |||
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              |||||||||||  || || |||||||| |||||||| ||||| || |||||||| ||||||
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              |||||||||||||||| || ||||| |||||||||| || ||||||||||||||||||
```

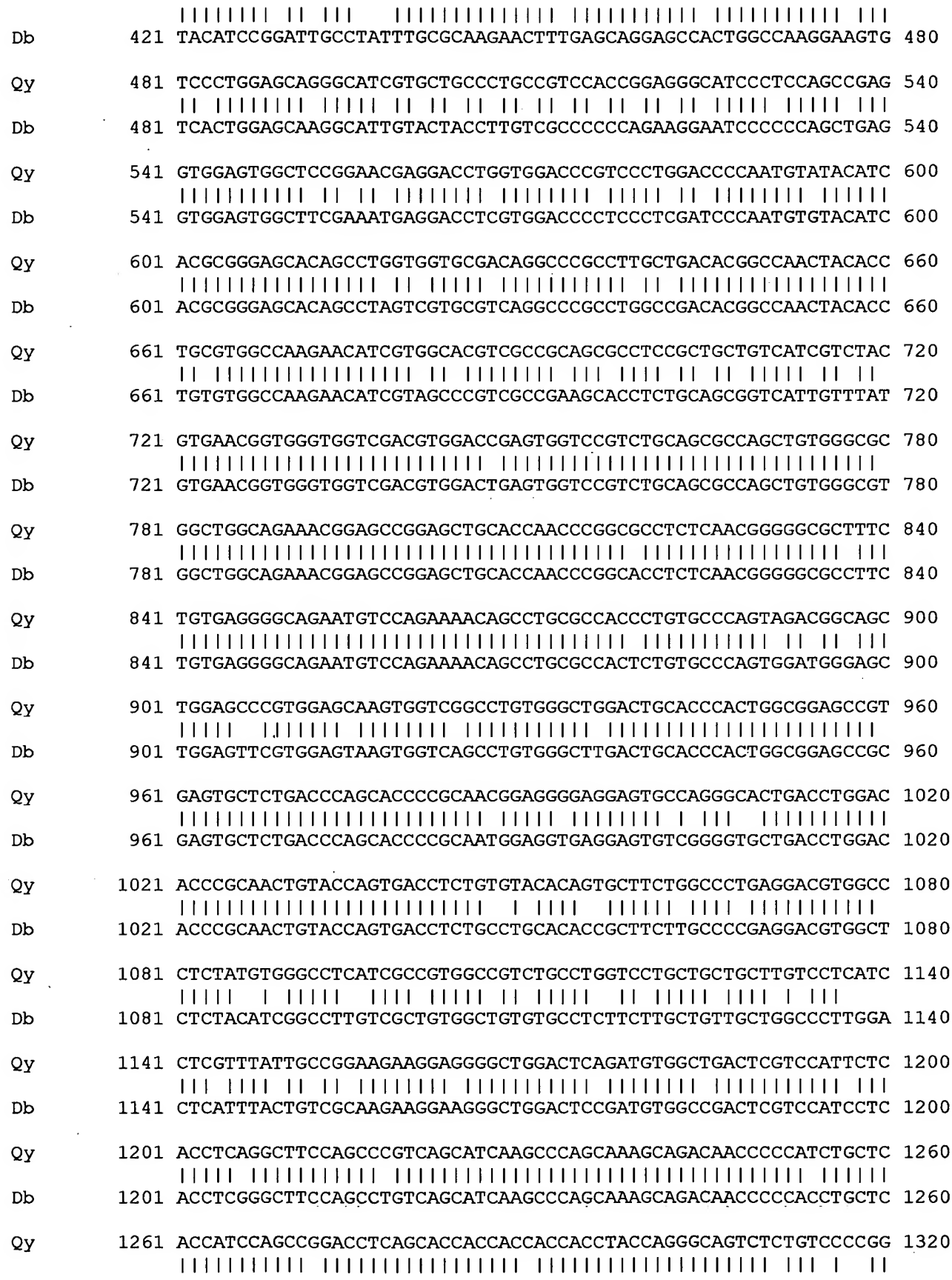




```
Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             |||||||| ||||||||||||||||||||  || ||||||||||||||||| |||||||
Db       2341 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
              |||| |||||||||||||| || ||||| ||||||||| | ||||| ||||||||||
Db       2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             |||||||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db       2521 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||| |||||||| |||||||||||||||||||||||| |||||||||| ||||| |||
Db       2581 AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             || |||||||||| ||||| |||||||||||||| |||||||||| ||||| |||
Db       2641 GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697
```

```
RESULT 8
RNU87305
LOCUS       RNU87305                2697 bp    mRNA    linear   ROD 15-MAY-1997
DEFINITION  Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.
ACCESSION   U87305
VERSION     U87305.1  GI:2055391
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K., Ackerman,S.L. and
            Tessier-Lavigne,M.
  TITLE     Vertebrate homologues of C. elegans UNC-5 are candidate netrin
            receptors
  JOURNAL   Nature 386 (6627), 833-838 (1997)
  MEDLINE   97271897
   PUBMED   9126742
REFERENCE   2  (bases 1 to 2697)
  AUTHORS   Leonardo,E.D., Hinck,L., Masu,M., Keino-Masu,K. and
            Tessier-Lavigne,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San
            Francisco, CA 94143-0452, USA
FEATURES             Location/Qualifiers
     source          1. .2697
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain and ventral spinal cord"
```

```
                     /dev_stage="18 day embyo and 13 day embryo"
     CDS             1. .2697
                     /codon_start=1
                     /product="transmembrane receptor UNC5H1"
                     /protein_id="AAB57678.1"
                     /db_xref="GI:2055392"
                     /translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN
                     YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN
                     GGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEEC
                     RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS
                     DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCSRQDGPSPKFQL
                     SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG
                     RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGV
                     LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC
                     YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV
                     QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT
                     QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEG
                     GVPALVGPSAFKIPFLIRQKIIASLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT
                     AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
ORIGIN

  Query Match             83.5%;  Score 2252.2;  DB 10;  Length 2697;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches 2419;  Conservative    0;  Mismatches  278;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||||| |||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || ||||||||||||||||||||||| |||||||| |||||||| ||||||||||
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| ||||||||||||||
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||| || | |||||||||||| ||||| ||||| ||||||||||||||||||| |||||
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||| |||||||| || ||||| || || |||||||| ||| | ||||| ||  ||||
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              |||||||||||||| || || || || |||||||||| ||||| ||||| |||||||||
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              |||||||||||||||| |||||||||||||||||||| |||||||| ||||||||||||
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
```



```
Db       1261 ACCATCCAGCCAGACCTCAGCACCACCACTACCACCTACCAGGGCAGTCTATGTTCGAGG 1320

Qy       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380
              |||||||| |||||||||||||||||||| | ||||| |||||||||||||| |||||
Db       1321 CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG 1380

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
               | |||||||| ||  |||||||||||| |||||||||||||| ||||| |||||||||
Db       1381 AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC 1440

Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
              |||||||||||||| |||||||| || |||||||||| |||||||||||||| ||||
Db       1441 CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC 1500

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
              |||||||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC 1560

Qy       1561 ATCCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
              || ||||| |||||||| |||||||| |||||||| |||||||||||||| |||||||||
Db       1561 ATACCCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG 1620

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
              || ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Db       1621 CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC 1680

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
              ||||| ||||| || |||||||||||||||||||||||| || ||||||||||||||
Db       1681 TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA 1740

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
              |||||||||||||||||||||| |||||||||||||||||||| |||||||||| |||
Db       1741 GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG 1800

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
              |||||||||||||||| || |||||| | || |||||||||||||||||||||||||||
Db       1801 GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG 1860

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
              ||| | |||||||| |||||||| |||||||||||||||||||||||| |||||||||
Db       1861 GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC 1920

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
              |||||||||||||| |||||||||| ||| |||||||||| || |||||||| ||
Db       1921 CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG 1980

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
              ||||| |||||||||||||| || |||||||| || |||||||||| || |||||||||
Db       1981 TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG 2040

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
              |||||||||||||||||||||| || |||||||||||||||||| || |||||||||
Db       2041 GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC 2100

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
              ||||| |||||||||||||||| || || |||||||||| ||||||||||||||||||
Db       2101 TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG 2160
```

```
Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||| ||||| |||||||||||||||||||||| ||||| |||||||||| |||||
Db       2161 AAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAACGGCACC 2220

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             |||| |||  |||||||||||||||||||||||  ||| | ||||||| || ||||||
Db       2221 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||| |||||||||||||||||||| || |||||||||||| ||||||||||
Db       2281 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             |||||||| ||||||||||||||||||  || |||||||||||||||| |||||||
Db       2341 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
              |||| |||||||||||||| || ||||| |||||||| | ||||| ||||||||||
Db       2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||
Db       2521 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||| |||||||| |||||||||||||||||||||| ||||||||| ||||| |||
Db       2581 AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             || |||||||||| ||||| |||||||||||||| |||||||||| ||||| |||
Db       2641 GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697
```

```
RESULT 9
BD057524
LOCUS       BD057524               3014 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Netrin receptors.
ACCESSION   BD057524
VERSION     BD057524.1  GI:22603130
KEYWORDS    JP 2001505062-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 3014)
  AUTHORS   Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
  TITLE     Netrin receptors
  JOURNAL   Patent: JP 2001505062-A 1 17-APR-2001;
            THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT     PN   JP 2001505062-A/1
            PD   17-APR-2001
            PF   19-FEB-1998 JP 1998536840
            PR   19-FEB-1997 US    08/808982
```

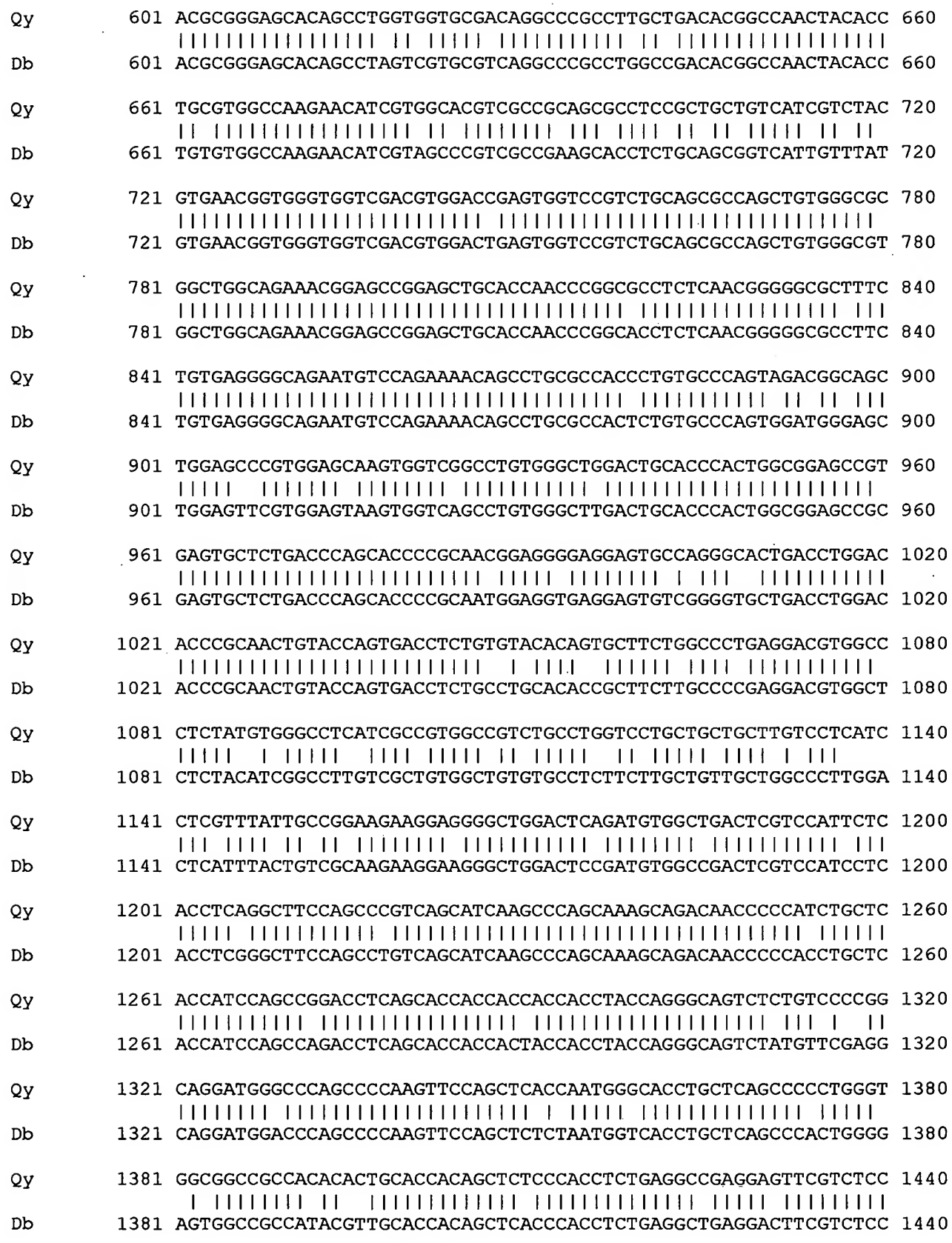
PI MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
PI MASU,
PI KAZUKO KEINO MASU
PC C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.

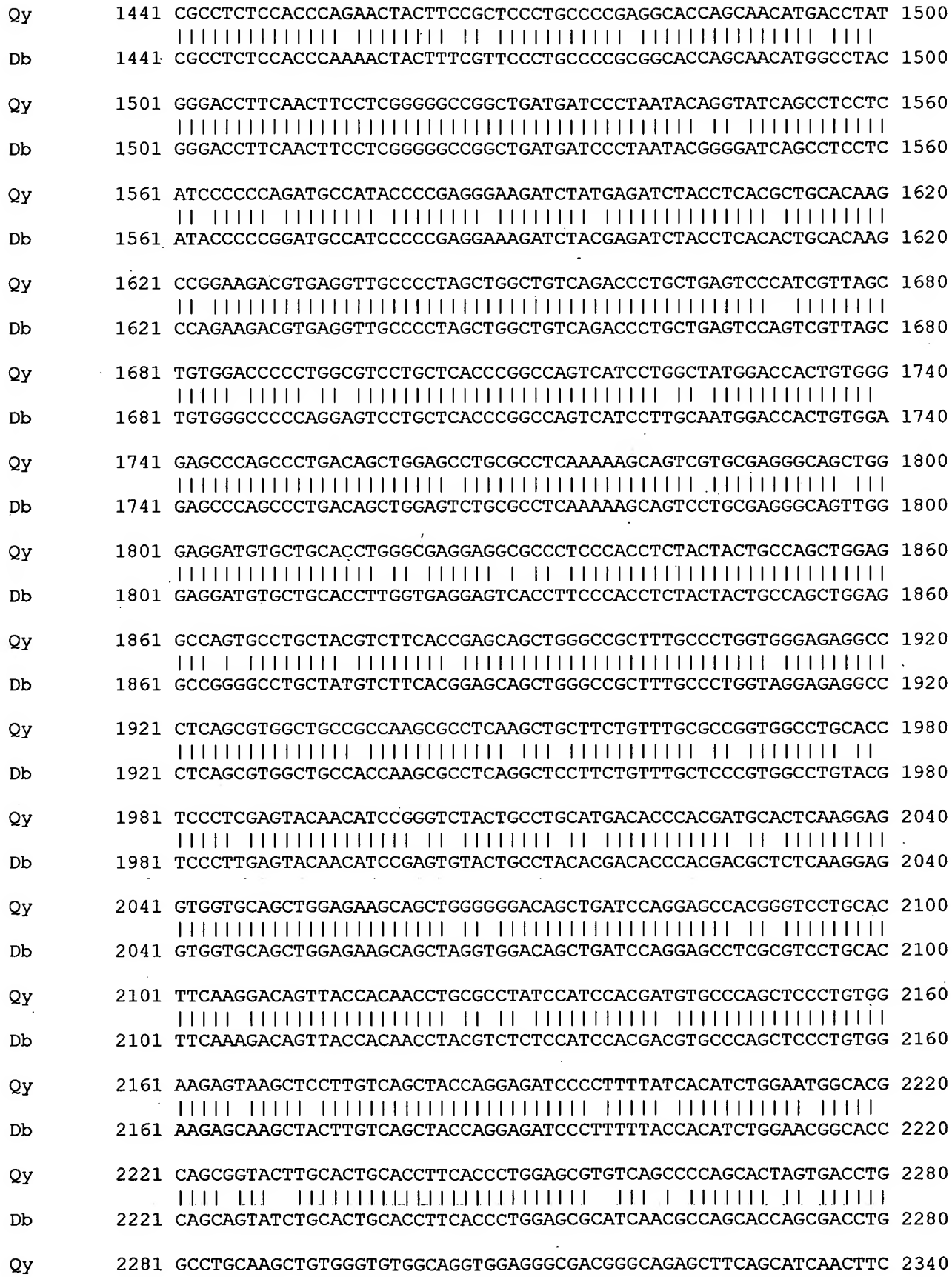
FEATURES Location/Qualifiers
source 1. .3014
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 83.5%; Score 2252.2; DB 6; Length 3014;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2419; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || |||||||||||||||||||||| |||||||| |||||||| ||||||||||
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| |||||||||||||
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||| || | |||||||||||| ||||| ||||| |||||||||||||||||||| |||||
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||| |||||||| || |||.|| || || |||||||| ||| | ||||| ||  ||||
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              |||||||||||||| || || || || |||||||||||| ||||| ||||| |||||||||
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              |||||||||||||||| |||||||||||||||||| |||||||| ||||||||||||
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||| || |||   |||||||||||||| |||||||||| |||||||||| |||
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              || |||||||| ||||| || || || || || || || || || ||||| ||||| |||
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              |||||||||| || || |||||||| |||||||| ||||| || |||||||| ||||||
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600
```





```
                ||||||||| |||||||||||||||||||||| || |||||||||||| ||||||||||
Db       2281 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
                ||||||||| ||||||||||||||||||||  || ||||||||||||||||| |||||||
Db       2341 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
                 |||| |||||||||||||| || ||||| ||||||||| | ||||| ||||||||||
Db       2461 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
                ||||||||||||||||| ||||||||||||||||||||||||| |||||||||||||||
Db       2521 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
                ||||| |||||||| |||||||||||||||||||||||| |||||||||| ||||| |||
Db       2581 AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
                || |||||||||||| ||||| |||||||||||||| |||||||||| ||||| |||
Db       2641 GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697
```

```
RESULT 10
BC058084
LOCUS       BC058084                3844 bp    mRNA    linear   ROD 21-OCT-2003
DEFINITION  Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone
            MGC:66671 IMAGE:6813463), complete cds.
ACCESSION   BC058084
VERSION     BC058084.1  GI:34784158
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3844)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
```

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 3844)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 126 Row: b Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23346570.

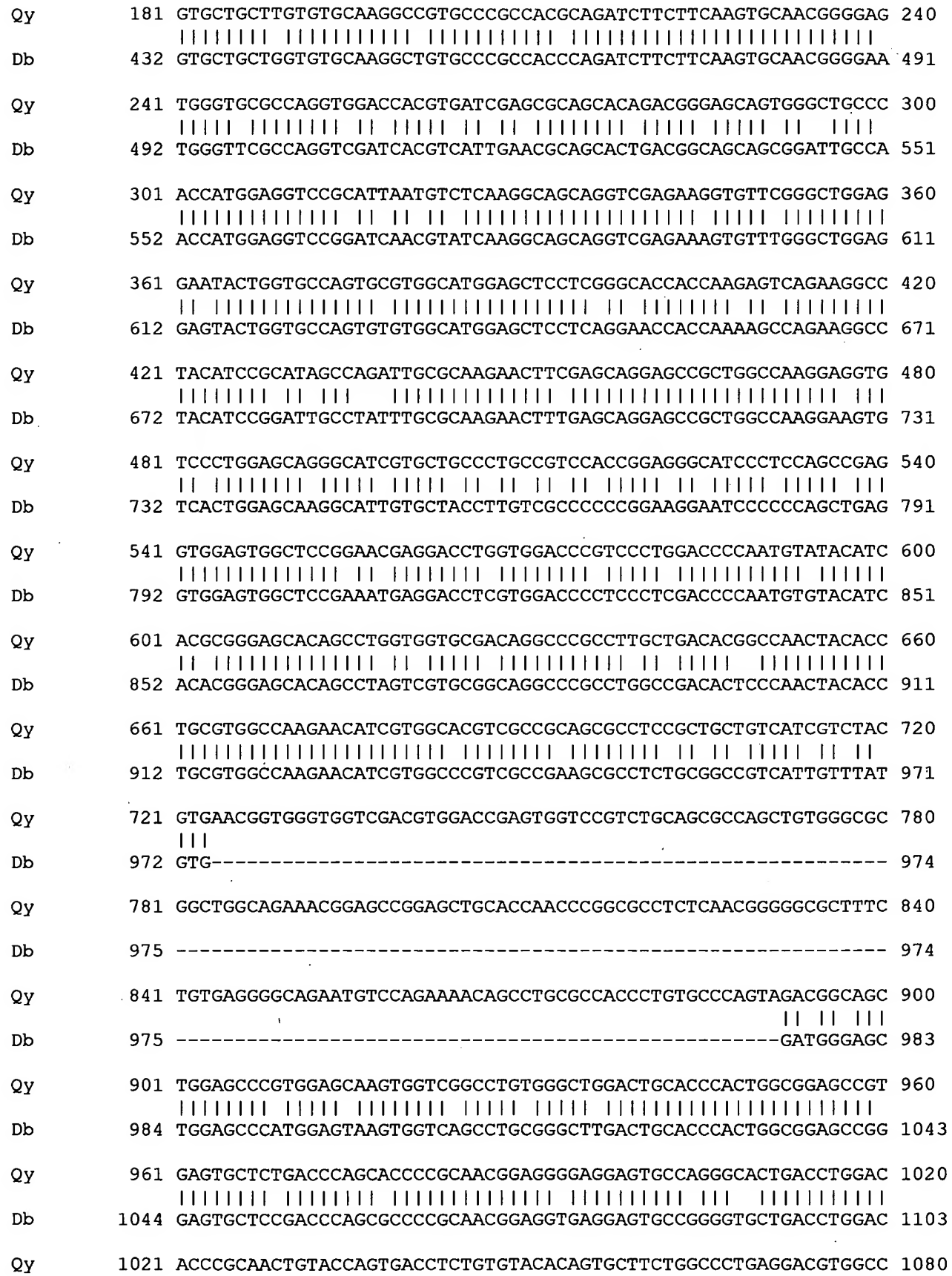
FEATURES Location/Qualifiers

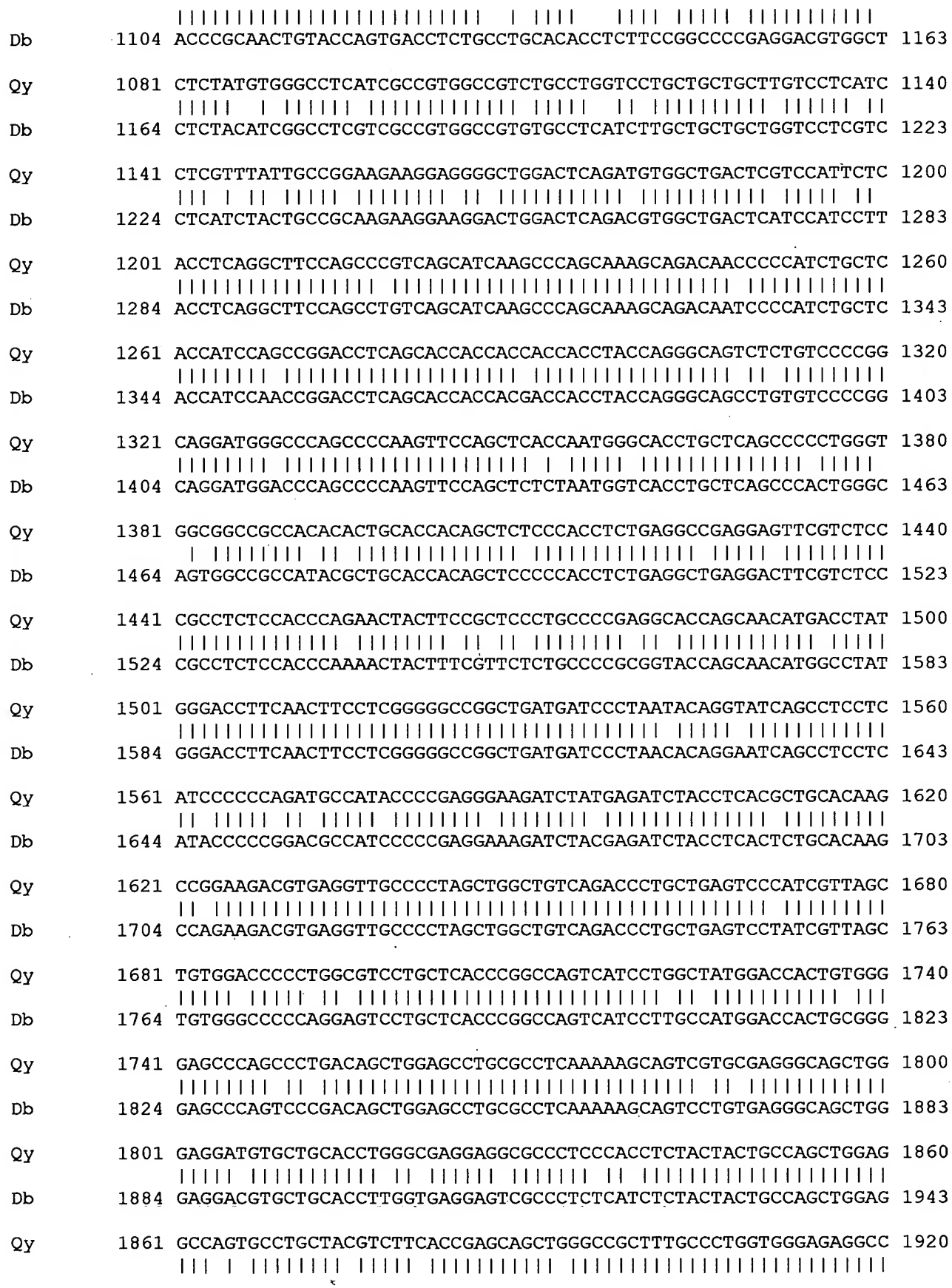
source 1. .3844
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:66671 IMAGE:6813463"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NIH_BMAP_FW0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

gene 1. .3844
/gene="Unc5a"
/db_xref="LocusID:107448"
/db_xref="MGI:894682"

CDS 252. .2780
/gene="Unc5a"
/codon_start=1
/product="Unc5a protein"
/protein_id="AAH58084.1"

```
                     /db_xref="GI:34784159"
                     /db_xref="LocusID:107448"
                     /db_xref="MGI:894682"
                     /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                     HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                     EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                     SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
                     YTCVAKNIVARRRSASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGE
                     ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
                     DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKF
                     QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
                     GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
                     GVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
                     ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
                     VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWN
                     GTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALES
                     EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPS
                     PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    714. .926
                     /gene="Unc5a"
                     /note="IG; Region: Immunoglobulin"
                     /db_xref="CDD:smart00409"
     misc_feature    984. .1130
                     /gene="Unc5a"
                     /note="TSP1; Region: Thrombospondin type 1 repeats"
                     /db_xref="CDD:smart00209"
     misc_feature    1566. .1877
                     /gene="Unc5a"
                     /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                     netrin receptors"
                     /db_xref="CDD:smart00218"
     misc_feature    2496. .2684
                     /gene="Unc5a"
                     /note="DEATH; Region: DEATH domain, found in proteins
                     involved in cell death (apoptosis). Alpha-helical domain
                     present in a variety of proteins with apoptotic functions.
                     Some (but not all) of these domains form homotypic and
                     heterotypic dimers"
                     /db_xref="CDD:smart00005"
ORIGIN

  Query Match             72.6%;  Score 1957.4;  DB 10;  Length 3844;
  Best Local Similarity   84.6%;  Pred. No. 0;
  Matches 2283;  Conservative    0;  Mismatches  246;  Indels  168;  Gaps    1;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| | || |||||
Db        252 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTT 311

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        312 CGTGGTTCGGGTGCCCAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCG 371

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| |||||||||||||
Db        372 GACCTGCTGCCCCACTTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCC 431
```





```
Db       1944 GCCGGGGCCTGCTATGTCTTTACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCC 2003

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
              |||||||||||||||| |||||||||||| ||| |||||||||| || |||||||| ||
Db       2004 CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACG 2063

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
              |||||||||||||||||||| || |||||||||| |||||||||||||| |||||||||
Db       2064 TCCCTCGAGTACAACATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAAGGAG 2123

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
              |||||||||||||||||||||||||| |||||||||||||||||| || |||||||||
Db       2124 GTGGTGCAGCTGGAGAAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCAC 2183

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
              ||||| |||||||||||||||| || || |||||||||| ||||||||||||||||||
Db       2184 TTCAAAGACAGTTACCACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGG 2243

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
              ||||| |||||||||||||||||||||||||| ||||| ||||||||||||||||
Db       2244 AAGAGCAAGCTCCTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACT 2303

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
              |||| |||  |||||||||||||||||||||| ||||   ||||||| || ||||||
Db       2304 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTG 2363

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
              ||||||||| |||||||||||||||||||| || || |||||||||| |||||||||
Db       2364 GCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTT 2423

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
              |||||||| |||||||| ||||||||||  |||||||||||||||||||| |||||||
Db       2424 AACATCACTAAGGACACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTC 2483

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
Db       2484 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATT 2543

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
               ||||||||||| |||||||| || ||||| |||||||||| ||||| ||||||||||
Db       2544 ACCAGCCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTT 2603

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||
Db       2604 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2663

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
              ||||| |||||||||||||||||||||||||||||||| |||||||||| ||||| |||
Db       2664 AACCTATGGGAGGCGCGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTG 2723

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
              || |||||||||||||||| |||||||||||||| ||||| ||||| |||||||||
Db       2724 GCCGGACTGGGCCAGCCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGA 2780
```

RESULT 11
BC009333

LOCUS BC009333 2688 bp mRNA linear PRI 25-MAR-2004
DEFINITION Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone IMAGE:4126760), partial cds.
ACCESSION BC009333
VERSION BC009333.2 GI:40226527
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2688)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2688)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT On Dec 19, 2003 this sequence version replaced gi:14424611.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,

McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 26 Row: g Column: 22.

```
FEATURES             Location/Qualifiers
     source          1. .2688
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4126760"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_19"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            <1. .2688
                     /gene="UNC5A"
                     /note="synonym: UNC5H1"
                     /db_xref="LocusID:90249"
                     /db_xref="MIM:607869"
     CDS             <1. .1627
                     /gene="UNC5A"
                     /codon_start=2
                     /product="UNC5A protein"
                     /protein_id="AAH09333.2"
                     /db_xref="GI:40226528"
                     /db_xref="LocusID:90249"
                     /db_xref="MIM:607869"
                     /translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT
                     SGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL
                     GGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGI
                     SLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA
                     MDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR
                     FALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQ
                     LIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT
                     LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS
                     AFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA
                     RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"
     misc_feature    413. .724
                     /gene="UNC5A"
                     /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                     netrin receptors"
                     /db_xref="CDD:smart00218"
     misc_feature    1343. .1594
                     /gene="UNC5A"
                     /note="DEATH; Region: DEATH domain, found in proteins
                     involved in cell death (apoptosis). Alpha-helical domain
                     present in a variety of proteins with apoptotic functions.
                     Some (but not all) of these domains form homotypic and
                     heterotypic dimers"
                     /db_xref="CDD:smart00005"
ORIGIN
```

Query Match 60.3%; Score 1625.4; DB 9; Length 2688;
Best Local Similarity 99.9%; Pred. No. 4.1e-269;

```
  Matches 1626;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1071 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCT 60

Qy       1131 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC 120

Qy       1191 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC 180

Qy       1251 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCT 240

Qy       1311 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG 300

Qy       1371 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA 360

Qy       1431 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA 420

Qy       1491 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT 1550
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        421 CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT 480

Qy       1551 CAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGCCTCCTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC 540

Qy       1611 GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCC 600

Qy       1671 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA 1730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA 660

Qy       1731 CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA 1790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA 720

Qy       1791 GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTG 1850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTG 780

Qy       1851 CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT 1910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT 840
```

```
Qy       1911 GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT 1970
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT 900

Qy       1971 GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGC 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGC 960

Qy       2031 ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG 2090
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACG 1020

Qy       2091 GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAG 2150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAG 1080

Qy       2151 CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG 2210
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG 1140

Qy       2211 GAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC 2270
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC 1200

Qy       2271 TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG 2330
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG 1260

Qy       2331 CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA 2390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA 1320

Qy       2391 AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA 2450
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA 1380

Qy       2451 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 2510
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 1440

Qy       2511 CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 2570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 1500

Qy       2571 CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGC 2630
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGC 1560

Qy       2631 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA 2690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA 1620

Qy       2691 GTGCTGA 2697
             |||||||
Db       1621 GTGCTGA 1627
```

RESULT 12
BD057525
LOCUS BD057525 1787 bp DNA linear PAT 27-AUG-2002
DEFINITION Netrin receptors.
ACCESSION BD057525
VERSION BD057525.1 GI:22603131
KEYWORDS JP 2001505062-A/2.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1787)
AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
TITLE Netrin receptors
JOURNAL Patent: JP 2001505062-A 2 17-APR-2001;
THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT PN JP 2001505062-A/2
PD 17-APR-2001
PF 19-FEB-1998 JP 1998536840
PR 19-FEB-1997 US 08/808982
PI MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
PI MASU,
PI KAZUKO KEINO MASU
PC C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
source 1. .1787
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 57.6%; Score 1552.4; DB 6; Length 1787;
Best Local Similarity 98.5%; Pred. No. 1.5e-256;
Matches 1651; Conservative 0; Mismatches 16; Indels 9; Gaps 8;

```
Qy       1025 GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT 1084
             |||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db          1 GCAACTGTACCAGTGACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCT 59

Qy       1085 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 1144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCG 119

Qy       1145 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 1204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT 179

Qy       1205 CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 1263
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        180 CAGGCTTCCAGCCCGTCAGCATCTAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC 239

Qy       1264 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 1323
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 ATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAG 299
```

```
Qy       1324 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 1383
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC 359

Qy       1384 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 1443
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC 419

Qy       1444 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 1503
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG 479

Qy       1504 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC 1563
             |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        480 ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC 539

Qy       1564 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 1623
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG 599

Qy       1624 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 1683
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT 659

Qy       1684 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 1743
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG 719

Qy       1744 CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG 1803
             |||||||||||||||||||||||  |||||||||||||||||||||| |||||||||
Db        720 CCCAGCCCTGACAGCTGGAGCCTGGCCCTCAAAAAGCAGTCGTGCGAGGG-AGCTGGGAG 778

Qy       1804 GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCC 1863
             |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GATGT-CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCC 837

Qy       1864 AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC 1923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC 897

Qy       1924 AGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC 1983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 AGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC 957

Qy       1984 CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG 2043
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG 1017

Qy       2044 GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTC 2103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTT- 1076

Qy       2104 AAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAG 2163
             |||||||||||||||||||  |||    ||||||||||||||||||||||||||||||||
Db       1077 AAGGACAGTTACCACAACCT--GCCCTATCATCCACGATGTGCCCAGCTCCCTGTGGAAG 1134
```

```
Qy       2164 AGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 2223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 AGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG 1194

Qy       2224 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCC 2283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCC 1254

Qy       2284 TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 2343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC 1314

Qy       2344 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 2403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA 1374

Qy       2404 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 2463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC 1434

Qy       2464 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC 2523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC 1494

Qy       2524 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC 2583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC 1554

Qy       2584 CTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCT 2643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1555 CTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCT 1614

Qy       2644 GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGA 2697
             ||    |||||||  |||| |||| || |||||||| ||||||||||||||||||
Db       1615 GGGACTGGCCAGCAGGACGGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGCTGA 1670
```

```
RESULT 13
AX054976
LOCUS       AX054976               9700 bp    DNA     linear   PAT 13-JAN-2001
DEFINITION  Sequence 91 from Patent WO0073328.
ACCESSION   AX054976
VERSION     AX054976.1  GI:12228344
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   van Criekinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
  TITLE     Unc-5 constructs and screening methods
  JOURNAL   Patent: WO 0073328-A 91 07-DEC-2000;
            Devgen NV (BE)
FEATURES             Location/Qualifiers
     source          1. .9700
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
```

/note="plasmid pGC1037"

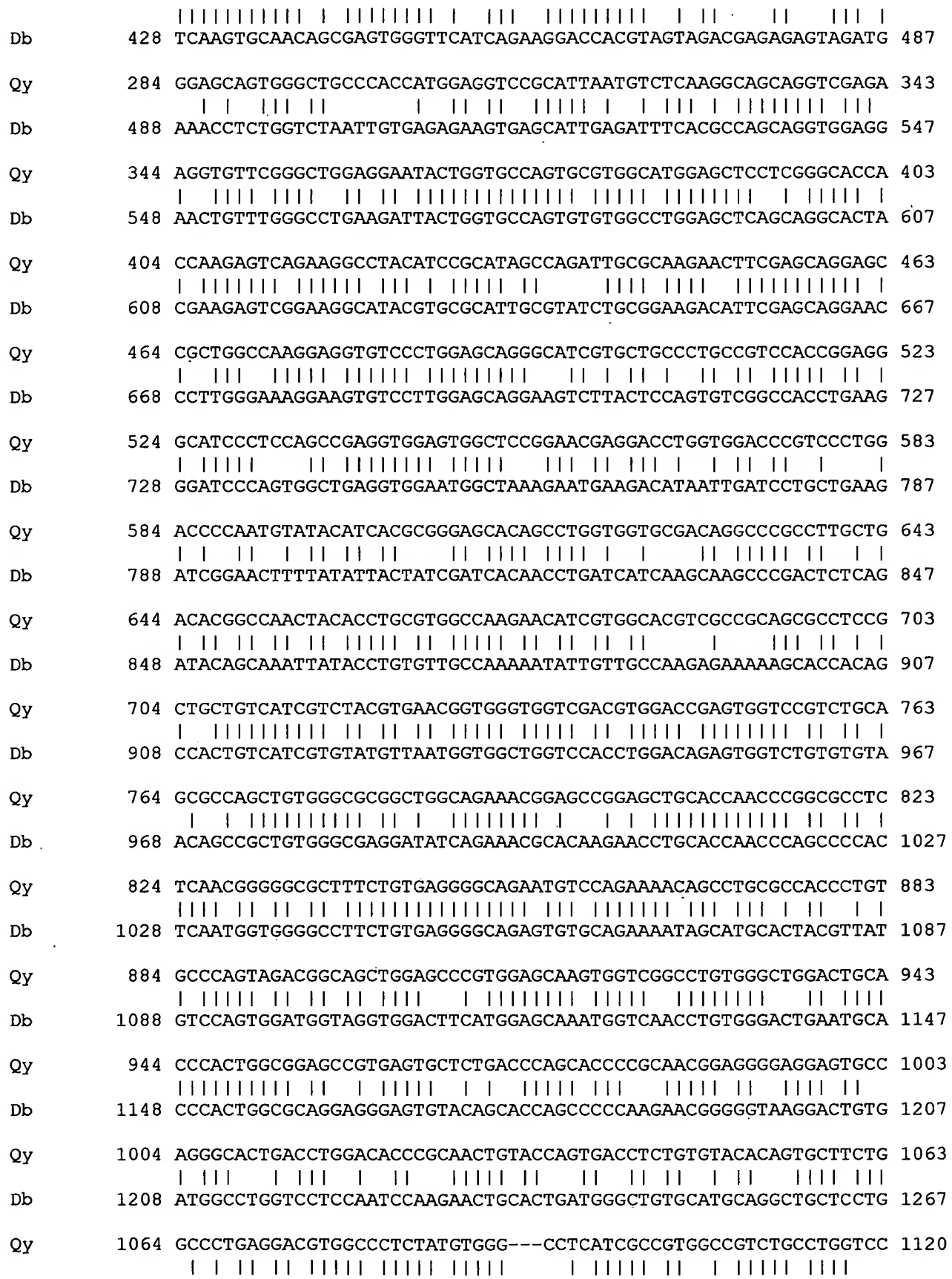
ORIGIN

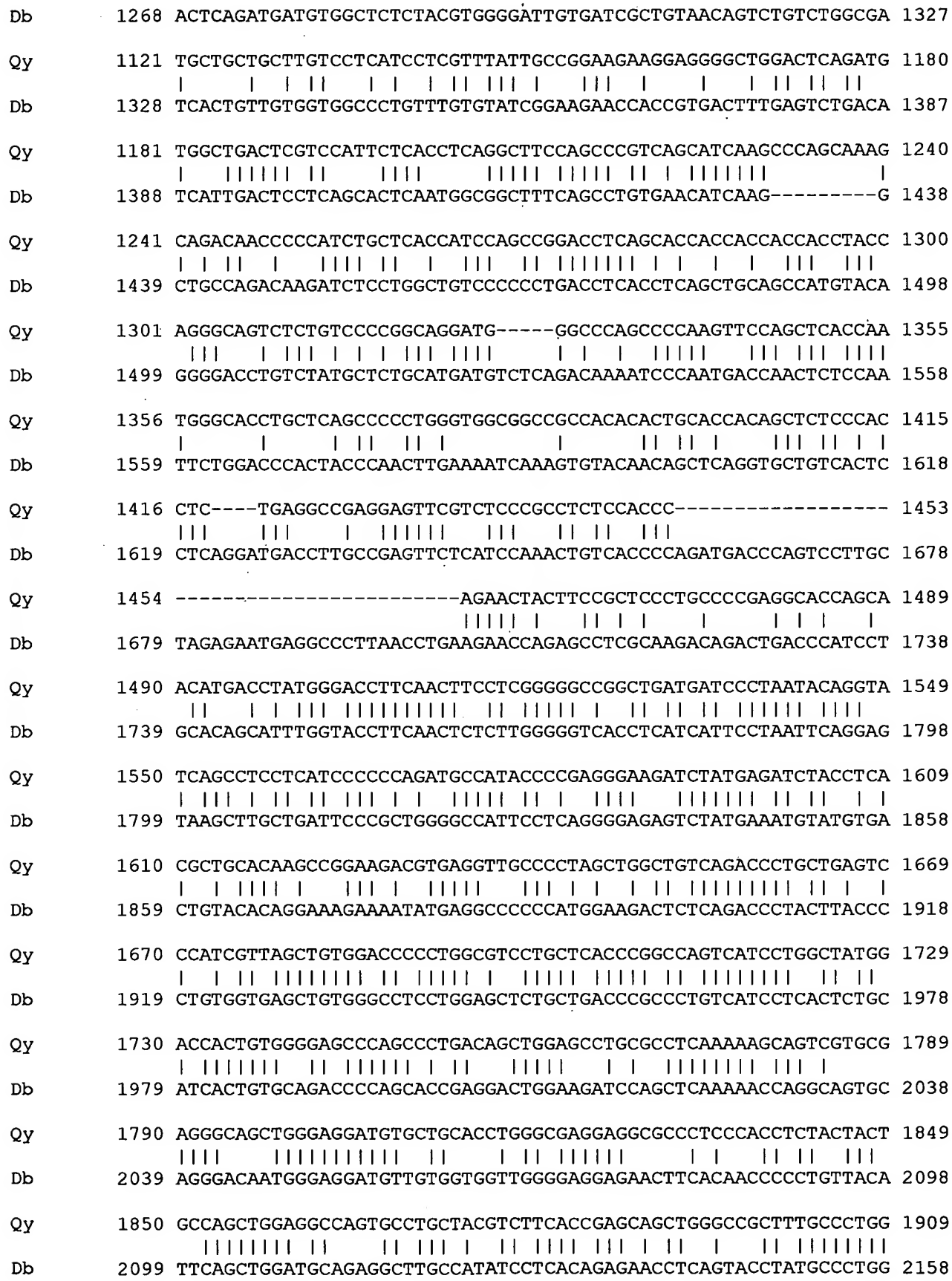
Query Match 48.3%; Score 1302.8; DB 6; Length 9700;
Best Local Similarity 99.8%; Pred. No. 8.4e-214;
Matches 1304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

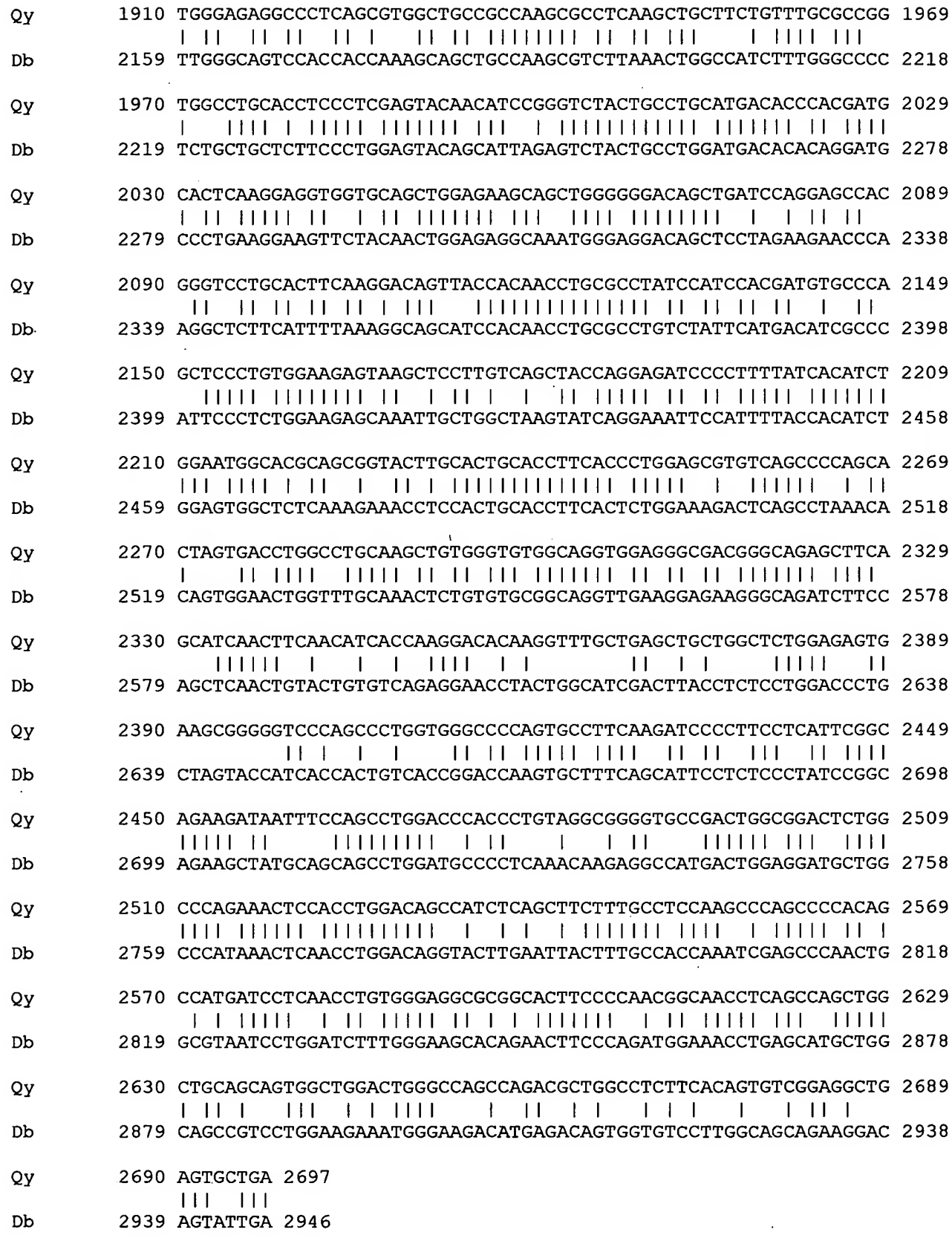
```
Qy       1392 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 66

Qy       1452 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 126

Qy       1512 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGA 1571
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        127 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCCAGA 186

Qy       1572 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 1631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 246

Qy       1632 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306

Qy       1692 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 1751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 366

Qy       1752 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT 1811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT 426

Qy       1812 GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG 1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG 486

Qy       1872 CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC 1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC 546

Qy       1932 TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA 1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTCGAGTA 606

Qy       1992 CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT 2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT 666

Qy       2052 GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG 2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG 726

Qy       2112 TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 2171
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        727 TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 786

Qy       2172 CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 846

Qy       2232 GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 906

Qy       2292 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 966

Qy       2352 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2411
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        967 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCTGGT 1026

Qy       2412 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 1086

Qy       2472 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 CCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 1146

Qy       2532 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 2591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGA 1206

Qy       2592 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 2651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 GGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGG 1266

Qy       2652 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
              |||||||||||||||||||||||||||||||||||||||||||||
Db       1267 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 1312


RESULT 14
MMU72634
LOCUS       MMU72634                9299 bp    mRNA    linear   ROD 13-MAY-1997
DEFINITION  Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
            complete cds.
ACCESSION   U72634
VERSION     U72634.1  GI:2088526
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 9299)
  AUTHORS   Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B.
            and Knowles,B.B.
  TITLE     The mouse rostral cerebellar malformation gene encodes an
            UNC-5-like protein
  JOURNAL   Nature 386 (6627), 838-842 (1997)
```

```
  MEDLINE   97271898
   PUBMED   9126743
REFERENCE   2  (bases 1 to 9299)
  AUTHORS   Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
            Bar Harbor, ME 04609, USA
FEATURES             Location/Qualifiers
     source          1. .9299
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57B6/SJL"
                     /db_xref="taxon:10090"
                     /chromosome="3"
                     /map="between Amy1 and Mpmv9"
     gene            1. .9299
                     /gene="rcm"
     CDS             151. .2946
                     /gene="rcm"
                     /function="transmembrane receptor"
                     /note="similar to C. elegans UNC-5"
                     /codon_start=1
                     /product="rostral cerebellar malformation protein"
                     /protein_id="AAB54103.1"
                     /db_xref="GI:2088527"
                     /translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                     WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA
                     ITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
                     GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS
                     LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE
                     MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHCADPSTEDWKIQL
                     KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL
                     KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN
                     LRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC
                     VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL
                     DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
                     EMGRHETVVSLAAEGQY"
ORIGIN

  Query Match             36.8%;  Score 992;  DB 10;  Length 9299;
  Best Local Similarity   62.7%;  Pred. No. 1.9e-160;
  Matches 1660;  Conservative    0;  Mismatches  925;  Indels   63;  Gaps    5;

Qy        104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163
              | ||   || |   || || |   | || ||||||||  | ||||||||||| |  ||||
Db        308 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 367

Qy        164 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 223
              | || ||||||||||| |||   ||    || || |||   || ||||| ||||||| ||
Db        368 TTGTGAAGAACAAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 427

Qy        224 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 283
```



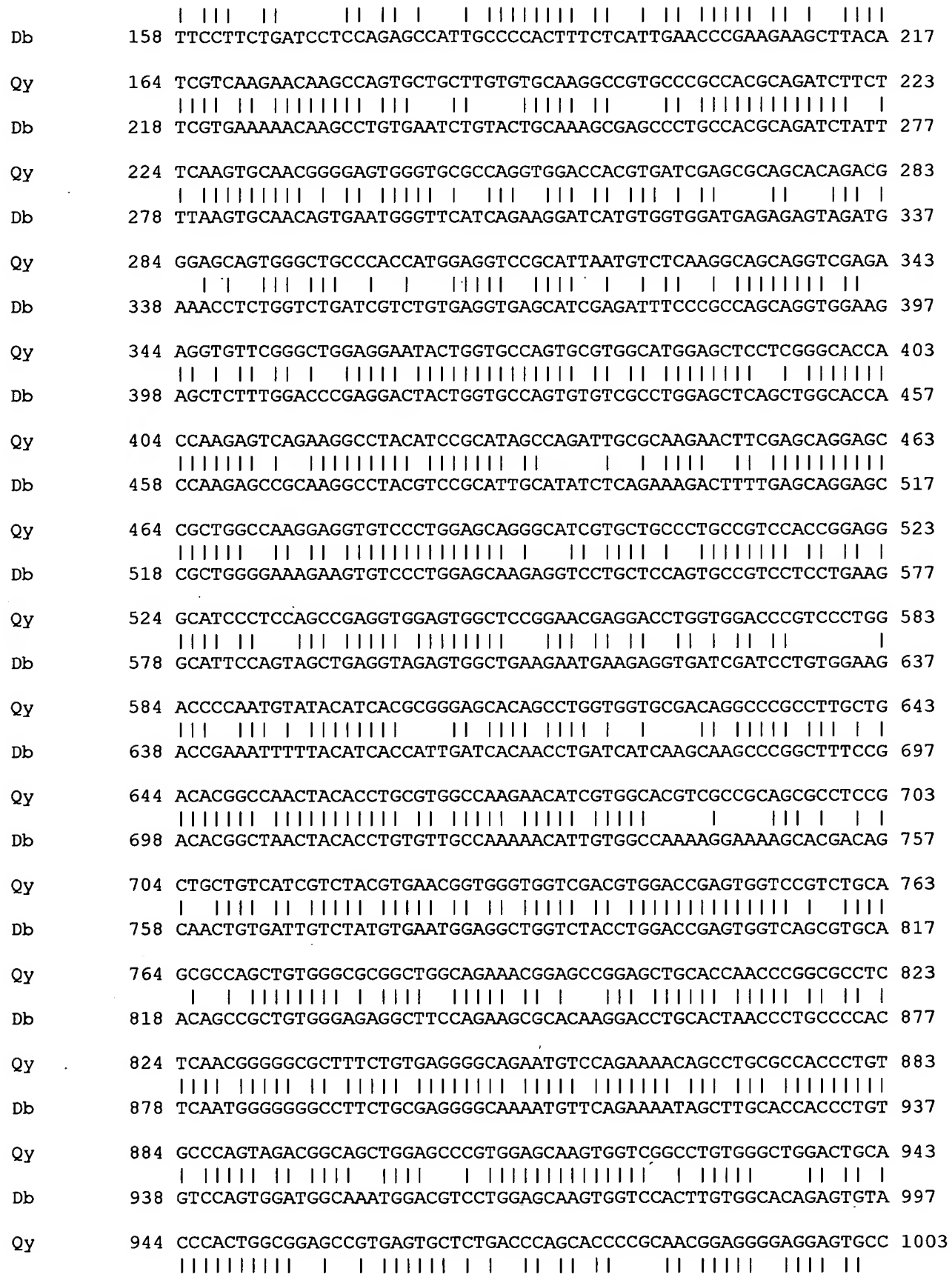


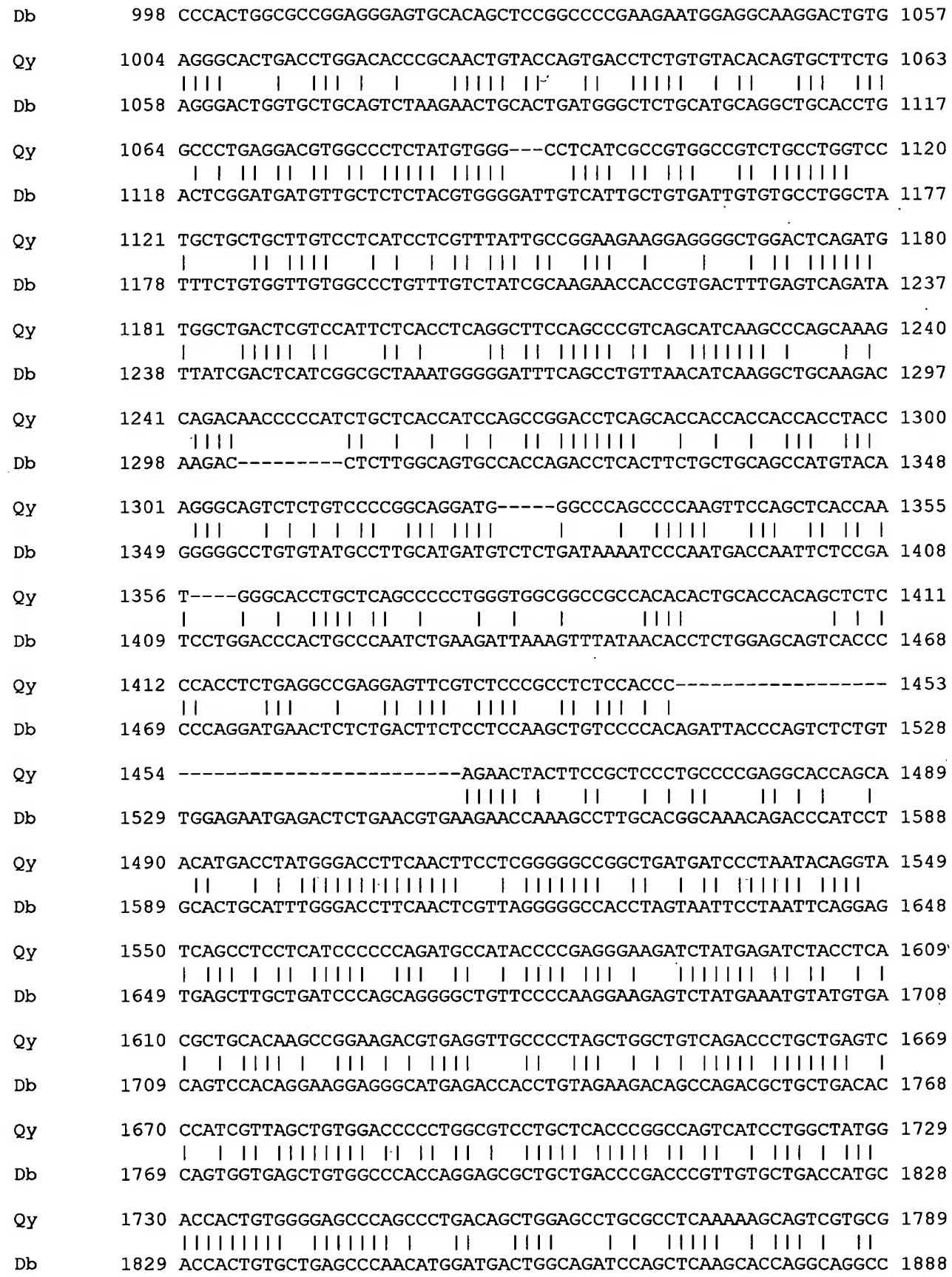


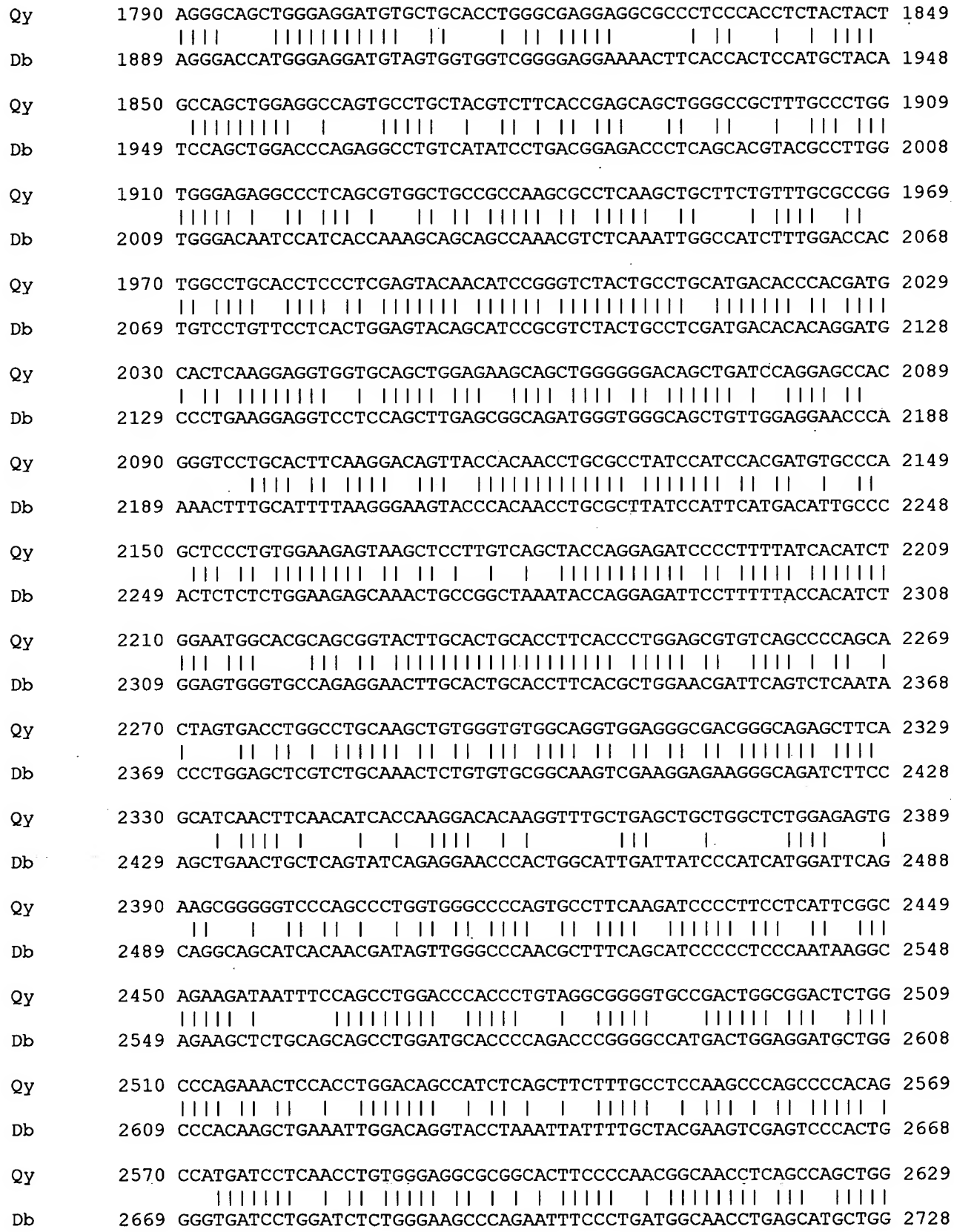
RESULT 15
AY187310
LOCUS AY187310 2962 bp mRNA linear VRT 06-JUN-2003
DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.
ACCESSION AY187310
VERSION AY187310.1 GI:31442350
KEYWORDS .
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2962)
AUTHORS Guan,W. and Condic,M.L.
TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development
JOURNAL Gene Expr. Patterns 3, 369-373 (2003)
REFERENCE 2 (bases 1 to 2962)
AUTHORS Guan,W. and Condic,M.L.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah, 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA
FEATURES Location/Qualifiers
source 1. .2962
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
CDS 1. .2796
/codon_start=1
/product="UNC5-like protein 3"
/protein_id="AAO67275.1"
/db_xref="GI:31442351"
/translation="MGKGLEGTAARCGLGMGYLLHSVVLPALAVLGASRPGSAAQDDD
FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
KDHVVDERVDETSGLIVCEVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEEVIDPVEDRNFYI
TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC
GRGFQKRTRTCTNPAPLNGGAFCEGQNVQKIACTTLCPVDGKWTSWSKWSTCGTECTH
WRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVIVCLA
ISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR
GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS
LLENETLNVKNQSLARQTDPSCTAFGTFNSLGGHLVIPNSGVSLLIPAGAVPQGRVYE
MYVTVHRKEGMRPPVEDSQTLLTPVVSCGPPGALLTRPVVLTMHHCAEPNMDDWQIQL
KHQAGQGPWEDVVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKAAAKRL
KLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKTLHFKGSTHN
LRLSIHDIAHSLWKSKLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKLC
VRQVEGEGQIFQLNCSVSEEPTGIDYPIMDSAGSITTIVGPNAFSIPLPIRQKLCSSL
DAPQTRGHDWRMLAHKLKLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE
EMGRHETVVSLAAEGNY"
ORIGIN

Query Match 36.7%; Score 988.8; DB 5; Length 2962;
Best Local Similarity 62.6%; Pred. No. 8e-160;
Matches 1658; Conservative 0; Mismatches 927; Indels 63; Gaps 5;

Qy 104 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 163







```
Qy       2630 CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG 2689
              | ||||   |||  | | |||| |  |  ||  | |   | |     |   | || |
Db       2729 CAGCAGTGCTGGAAGAAATGGGACGACATGAAACCGTTGTTTCTTTGGCAGCAGAAGGAA 2788

Qy       2690 AGTGCTGA 2697
              | | ||||
Db       2789 ATTACTGA 2796
```

Search completed: March 6, 2005, 05:25:12
Job time : 11476.2 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 14:42:51 ; Search time 1373 Seconds
(without alignments)
11628.216 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggcccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2697	100.0	2752	6	ABK37922	Abk37922 cDNA enco
2	2697	100.0	2752	12	ADH71617	Adh71617 Human gen
3	2687.4	99.6	2697	6	ABK52891	Abk52891 Human net
4	2621.4	97.2	2881	6	ABK49422	Abk49422 DNA encod

5	2621.4	97.2	2881	12	ADH71609	Adh71609 Human gen
6	2619.8	97.1	2880	12	ADH71633	Adh71633 Human gen
7	2619.8	97.1	2881	12	ADH71649	Adh71649 Human gen
8	2619.8	97.1	2881	12	ADH71635	Adh71635 Human gen
9	2619.8	97.1	2881	12	ADH71637	Adh71637 Human gen
10	2619.8	97.1	2881	12	ADH71641	Adh71641 Human gen
11	2619.8	97.1	2881	12	ADH71629	Adh71629 Human gen
12	2619.8	97.1	2881	12	ADH71631	Adh71631 Human gen
13	2619.8	97.1	2881	12	ADH71645	Adh71645 Human gen
14	2619.8	97.1	2881	12	ADH71627	Adh71627 Human gen
15	2619.8	97.1	2881	12	ADH71639	Adh71639 Human gen
16	2619.8	97.1	2881	12	ADH71643	Adh71643 Human gen
17	2619.8	97.1	2881	12	ADH71625	Adh71625 Human gen
18	2619.8	97.1	2881	12	ADH71647	Adh71647 Human gen
19	2618.2	97.1	2881	10	ADG42568	Adg42568 Novel hum
20	2368.6	87.8	2907	4	AAK52261	Aak52261 Human pol
21	2343	86.9	3561	12	ADL06497	Adl06497 Human tum
22	2343	86.9	3580	6	ABK15169	Abk15169 Human REP
23	2318	85.9	2575	12	ADH71621	Adh71621 Human gen
24	2266.2	84.0	2463	12	ADH71623	Adh71623 Human gen
25	2252.2	83.5	2697	6	AAS16843	Aas16843 Rat netri
26	2252.2	83.5	3014	2	AAV52940	Aav52940 Rat UNC-5
27	1625.4	60.3	2635	11	ADN95100	Adn95100 Human LEC
28	1552.4	57.6	1787	2	AAV52941	Aav52941 Human UNC
29	1302.8	48.3	9700	4	AAC90958	Aac90958 Plasmid p
30	1200.6	44.5	1321	4	AAH99530	Aah99530 Human pro
31	987.6	36.6	1002	12	ADH71619	Adh71619 Human gen
32	985.6	36.5	1009	12	ADH71613	Adh71613 Human gen
33	970	36.0	3646	5	AAS75738	Aas75738 DNA encod
34	970	36.0	3646	13	ADR99257	Adr99257 Human unc
35	968.8	35.9	2796	10	AAL56266	Aal56266 Human thr
36	943.4	35.0	993	12	ADH71611	Adh71611 Human gen
37	936.2	34.7	2860	6	ABT06279	Abt06279 Human NOV
38	936.2	34.7	2860	6	ABT06280	Abt06280 Human NOV
39	912.2	33.8	2895	6	ABQ93898	Abq93898 Human tra
40	904.2	33.5	2986	13	ADR99249	Adr99249 Human 1RO
41	904.2	33.5	3933	13	ADR07892	Adr07892 Full leng
42	901	33.4	3884	4	AAS21316	Aas21316 Human cDN
43	901	33.4	3884	8	ACA03675	Aca03675 cDNA enco
44	901	33.4	3884	8	ABX89213	Abx89213 DNA encod
45	901	33.4	3884	8	ACD41867	Acd41867 Human sec

ALIGNMENTS

```
RESULT 1
ABK37922
ID   ABK37922 standard; cDNA; 2752 BP.
XX
AC   ABK37922;
XX
DT   21-MAY-2002  (first entry)
XX
DE   cDNA encoding Human protein NOV1.
XX
KW   Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;
```

KW cell signal processing disorder; metabolic disorder; obesity; infection;
KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW psychosis; neurological disorder; anxiety; schizophrenia;
KW manic depression; dementia; dyskinesia; Huntington's disease;
KW Gilles de la Tourette's syndrome; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200210216-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US024225.
XX
PR 28-JUL-2000; 2000US-0221409P.
PR 04-AUG-2000; 2000US-0222840P.
PR 04-AUG-2000; 2000US-0223752P.
PR 04-AUG-2000; 2000US-0223762P.
PR 04-AUG-2000; 2000US-0223769P.
PR 04-AUG-2000; 2000US-0223770P.
PR 14-AUG-2000; 2000US-0225146P.
PR 15-AUG-2000; 2000US-0225392P.
PR 15-AUG-2000; 2000US-0225470P.
PR 16-AUG-2000; 2000US-0225697P.
PR 01-FEB-2001; 2001US-0263662P.
PR 05-APR-2001; 2001US-0281645P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;
XX
DR WPI; 2002-180074/23.
DR P-PSDB; AAU85403.
XX
PT New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT immune disorders.
XX
PS Claim 9; Page 9-10; 213pp; English.
XX
CC The invention relates to an isolated cytoplasmic, nuclear, membrane
CC bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC form. Also included are the nucleic acids encoding the NOVX proteins, a
CC vector comprising the nucleic acid, a cell comprising the vector, an anti
CC -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC antibody are useful for treating or preventing a NOVX-associated
CC disorder, where the disorder is selected from cardiomyopathy,
CC atherosclerosis, diabetes, a disorder related to cell signal processing
CC and metabolic pathway modulation, metabolic disorders, obesity,
CC infectious disease, anorexia, cancer-associated cachexia, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,

CC immune disorders, haematopoietic disorders, and the various
CC dyslipidaemias, metabolic disturbances associated with obesity, the
CC metabolic syndrome X and wasting disorders associated with chronic
CC diseases, bacterial, fungal, protozoal and viral infections, pain,
CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC hypertrophy, and psychotic and neurological disorders, including anxiety,
CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,
CC such as Huntington's disease and Gilles de la Tourette's syndrome. The
CC nucleic acid is useful in gene therapy. The present sequence encodes a
CC NOVX protein
XX
SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2697; DB 6; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885

Qy        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 945

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 1005

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1065

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1125

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1185

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1245

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1305

Qy       1261 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1365

Qy       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1425

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1485
```

```
Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1545

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1605

Qy       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1665

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1725

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1785

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1845

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1905

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1965

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2025

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2085

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2145

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2325
```

```
Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
```

```
RESULT 2
ADH71617
ID   ADH71617 standard; DNA; 2752 BP.
XX
AC   ADH71617;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Human gene of the invention NOV21e SEQ ID NO:513.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
XX
PF   03-JUN-2003; 2003WO-US017430.
XX
PR   03-JUN-2002; 2002US-0385120P.
PR   04-JUN-2002; 2002US-0385784P.
PR   05-JUN-2002; 2002US-0386041P.
PR   05-JUN-2002; 2002US-0386047P.
```

PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.

PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71618.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 513; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

Query Match 100.0%; Score 2697; DB 12; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;

```
  Matches 2697;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 645

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 705

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 765

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 825

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 885
```

```
Qy        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 945

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 1005

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1065

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1125

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1185

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1245

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1305

Qy       1261 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 ACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG 1365

Qy       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1425

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1485

Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1545

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1605

Qy       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1665

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1725
```

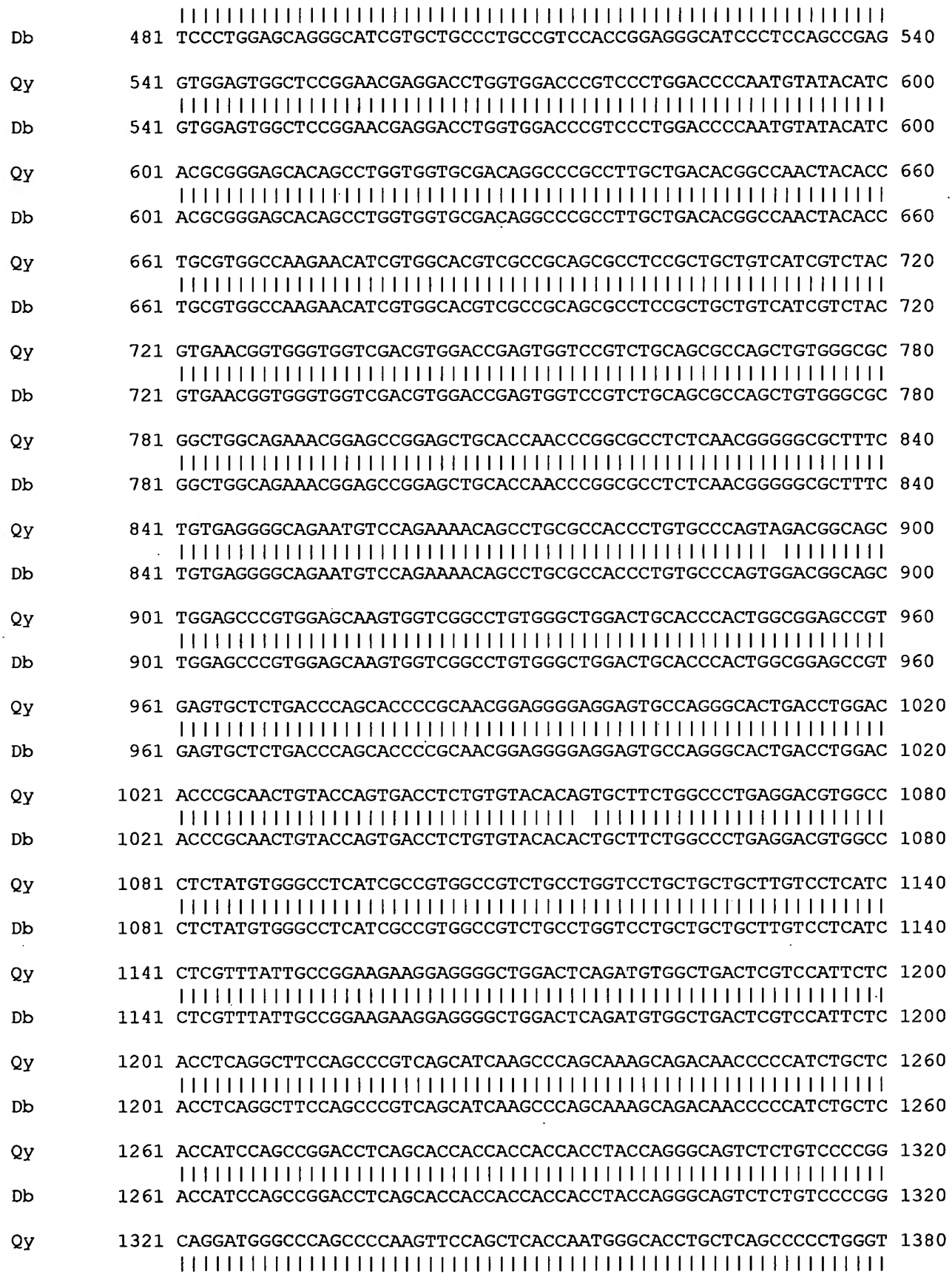
```
Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1726 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1785

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1786 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1845

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1905

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1965

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 2025

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2026 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2085

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2086 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2145

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2206 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265

Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2325

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2326 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2446 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2626 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2685

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742


RESULT 3
ABK52891
ID   ABK52891 standard; DNA; 2697 BP.
XX
AC   ABK52891;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW   Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW   nootropic; neuroprotective; cytostatic; antiparkinsonian;
KW   cerebroprotective; cancer; central nervous system; CNS; stroke;
KW   Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2697
FT                   /*tag=  a
FT                   /product= "Netrin binding membrane receptor UNC5H-1"
XX
PN   WO200233080-A2.
XX
PD   25-APR-2002.
XX
PF   15-OCT-2001; 2001WO-EP011891.
XX
PR   16-OCT-2000; 2000US-0240061P.
XX
PA   (FARB ) BAYER AG.
XX
PI   Koehler RH;
XX
DR   WPI; 2002-463314/49.
DR   P-PSDB; AAU97899.
XX
PT   Novel human netrin binding membrane receptor polypeptide and
PT   polynucleotides for identifying modulating agents useful in treating
PT   diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
PT   Alzheimer's disease.
XX
PS   Claim 1; Fig 1; 94pp; English.
XX
CC   This invention relates to the DNA and protein sequences of a novel
```

CC purified human netrin binding membrane receptor, UNC5H-1. The DNA
CC sequence of the invention is useful as a probe for detecting a nucleic
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences
CC of the invention are useful to screen for agents which decrease the
CC activity of the UNC5H-1 protein. The sequences are also useful for
CC screening agents which regulate (modulate) the activity of the protein of
CC the invention. A pharmaceutical composition containing the protein of the
CC invention or a reagent that modulates the activity of the UNC5H-1 protein
CC may be useful for treating a UNC5H-1 dysfunction related disease such as
CC cancer or a central nervous system (CNS) disorders (e.g, Parkinson's
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
CC proteins comprising the UNC5H-1 protein are useful for generating
CC antibodies and for in various assay systems, and the protein can be used
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method
CC of the invention is useful for detecting a coding sequence for the UNC5H-
CC 1 protein. The present sequence represents a DNA sequence encoding the
CC human netrin binding membrane receptor UNC5H-1 protein of the invention
XX
SQ Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Query Match 99.6%; Score 2687.4; DB 6; Length 2697;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db        421 TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
```



```
Db       1321 CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT 1380

Qy       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC 1440

Qy       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT 1500

Qy       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC 1560
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1501 GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC 1560

Qy       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG 1620

Qy       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC 1680

Qy       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG 1740

Qy       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG 1800

Qy       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG 1860

Qy       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC 1920

Qy       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC 1980

Qy       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG 2040

Qy       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 2100

Qy       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160

Qy       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
```

```
Qy       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280

Qy       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340

Qy       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400

Qy       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460

Qy       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520

Qy       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580

Qy       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 AACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG 2640

Qy       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
```

```
RESULT 4
ABK49422
ID   ABK49422 standard; DNA; 2881 BP.
XX
AC   ABK49422;
XX
DT   15-JUL-2002  (first entry)
XX
DE   DNA encoding human UNC5-like protein NOV1.
XX
KW   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
KW   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW   chromosome 13; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             87. .2786
FT                   /*tag=  a
FT                   /product= "Human UNC5-like protein NOV1"
XX
PN   WO200229038-A2.
```

XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031377.
XX
PR 04-OCT-2000; 2000US-0237862P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Herrmann JL, Rastelli L, Shimkets RA;
XX
DR WPI; 2002-340104/37.
DR P-PSDB; AAU79939.
XX
PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT treating cardiomyopathy, artherosclerosis, and cancer.
XX
PS Claim 8; Page 7-8; 180pp; English.
XX
CC The present invention relates to a new NOVX polypeptide having a 900
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
CC residue amino acid sequence, as given in the specification. The novel
CC polypeptide, and its encoding polynucleotide, are used to treat
CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC signal processing and metabolic pathway modulation, in a human. Detecting
CC the polypeptide or polynucleotide is useful for identifying cancerous
CC tissue. The antibody can be used to treat diabetes or cancer. The host
CC cells can be used to produce non-human transgenic animals useful in drug
CC screening. The present nucleic acid sequence is that of the human UNC5-
CC like NOV1 gene located on chromosome 13. This sequence encodes the human
CC UNC5-like protein NOV1 of the invention
XX
SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2621.4; DB 6; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386
```

```
Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226
```

```
Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
             |||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy        2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy        2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy        2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy        2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy        2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy        2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy        2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy        2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy        2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy        2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy        2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy        2695 TGA 2697
                |||
Db        2784 TGA 2786


RESULT 5
ADH71609
ID   ADH71609 standard; DNA; 2881 BP.
XX
AC   ADH71609;
```

XX

DT 25-MAR-2004 (first entry)

XX

DE Human gene of the invention NOV21a SEQ ID NO:505.

XX

KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.

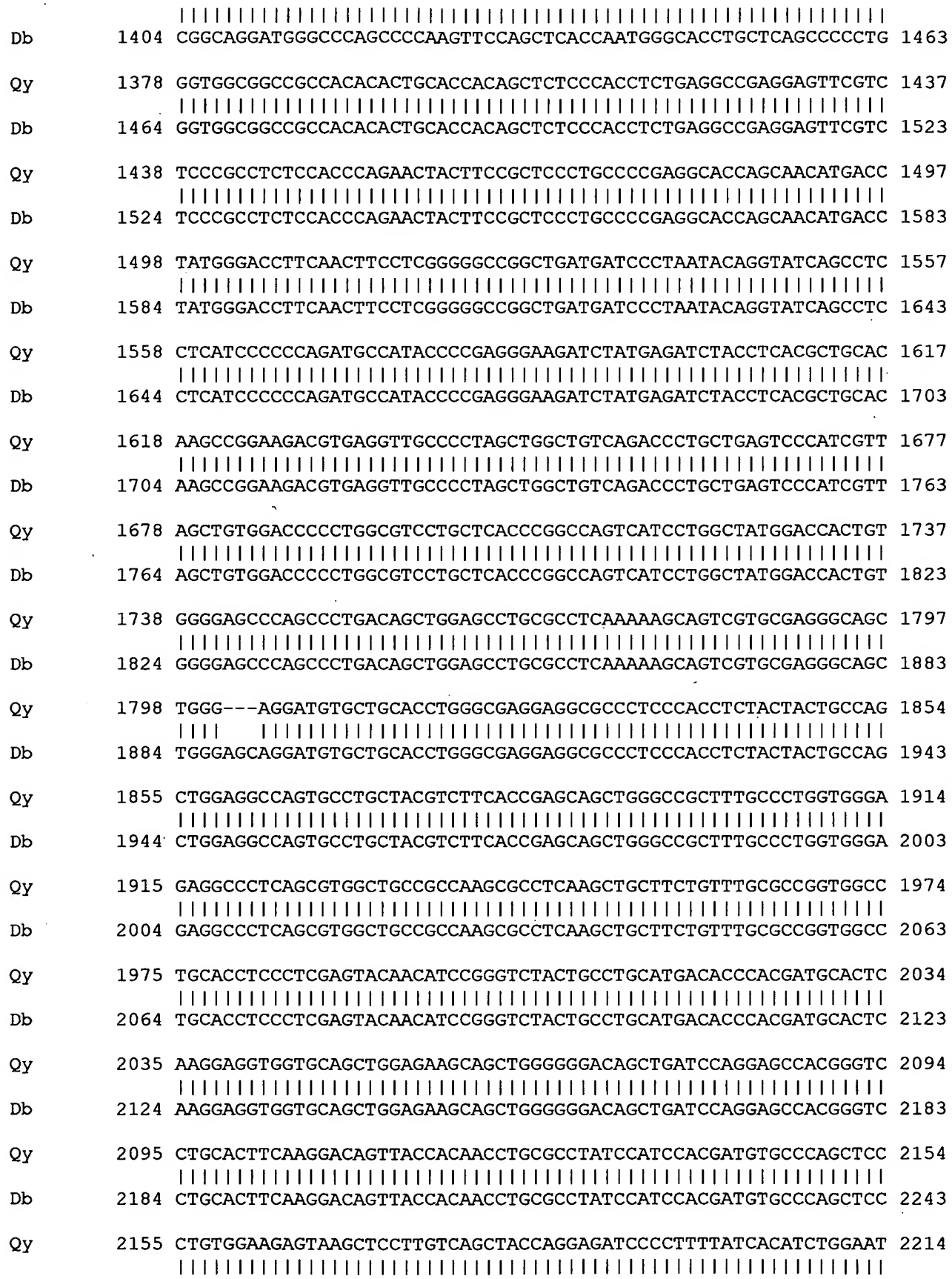
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71610.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX
PS Example 21; SEQ ID NO 505; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2621.4; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566
```

```
Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
```



```
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||
Db       2784 TGA 2786
```

RESULT 6
ADH71633
ID   ADH71633 standard; DNA; 2880 BP.
XX
AC   ADH71633;
XX
DT   25-MAR-2004 (first entry)
XX
DE   Human gene of the invention NOV21m SEQ ID NO:529.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX

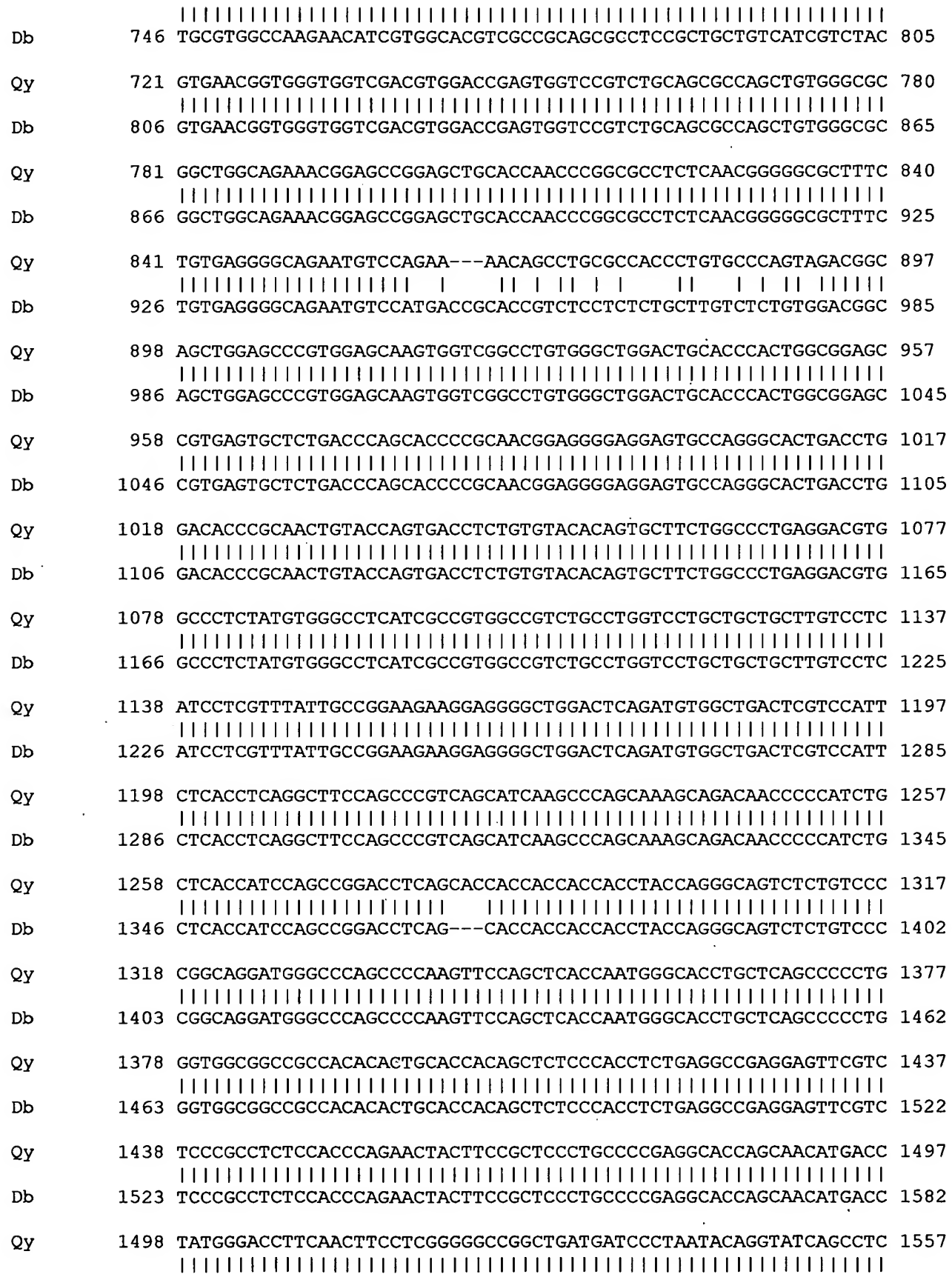
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.

PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71634.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 529; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,

CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2880 BP; 527 A; 984 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2880;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 145

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 205

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 265

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 325

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        326 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 385

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 445

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 505

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 565

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 625

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 685

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 745

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
```



```
Db       1583 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1642

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1643 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1702

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1703 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1762

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1763 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1822

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1823 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCACCTCAAAAAGCAGTCGTGCGAGGGCAGC 1882

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1883 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1942

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1943 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2002

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2003 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2062

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2063 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2122

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2123 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2182

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2183 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2242

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2243 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2302

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2303 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2362

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2363 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2422

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2423 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2482
```

```
Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2483 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2542

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2543 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2602

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2603 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2662

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2663 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2722

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2723 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2782

Qy       2695 TGA 2697
             |||
Db       2783 TGA 2785
```

RESULT 7
ADH71649
ID ADH71649 standard; DNA; 2881 BP.
XX
AC ADH71649;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV21u SEQ ID NO:545.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.

PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71650.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 545; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy     61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy    121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy    181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy    241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||   ||| 
Db    327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy    301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy    361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy    421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy    481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy    541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy    601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy    661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy    721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy    781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy    841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
          ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
```

```
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823
```

```
Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663
```

```
Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGCGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||   .
Db       2784 TGA 2786
```

RESULT 8
ADH71635
ID   ADH71635 standard; DNA; 2881 BP.
XX
AC   ADH71635;
XX
DT   25-MAR-2004 (first entry)
XX
DE   Human gene of the invention NOV21n SEQ ID NO:531.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
XX
PF   03-JUN-2003; 2003WO-US017430.
XX
PR   03-JUN-2002; 2002US-0385120P.
PR   04-JUN-2002; 2002US-0385784P.
PR   05-JUN-2002; 2002US-0386041P.
PR   05-JUN-2002; 2002US-0386047P.
PR   06-JUN-2002; 2002US-0386376P.
PR   06-JUN-2002; 2002US-0386453P.
PR   06-JUN-2002; 2002US-0386864P.
PR   06-JUN-2002; 2002US-0387016P.
PR   07-JUN-2002; 2002US-0386796P.
PR   07-JUN-2002; 2002US-0386816P.
PR   07-JUN-2002; 2002US-0386931P.
PR   07-JUN-2002; 2002US-0386942P.
PR   07-JUN-2002; 2002US-0386971P.
PR   07-JUN-2002; 2002US-0387262P.
PR   08-JUN-2002; 2002US-0296960P.
PR   10-JUN-2002; 2002US-0387400P.
PR   10-JUN-2002; 2002US-0387535P.
PR   11-JUN-2002; 2002US-0387610P.
PR   11-JUN-2002; 2002US-0387625P.
PR   11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX

PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71636.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 531; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166
```

```
Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
             ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACATCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003
```

```
Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
              |||
Db       2784 TGA 2786
```

RESULT 9
ADH71637
ID   ADH71637 standard; DNA; 2881 BP.
XX
AC   ADH71637;
XX
DT   25-MAR-2004 (first entry)
XX
DE   Human gene of the invention NOV21o SEQ ID NO:533.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
XX
PF   03-JUN-2003; 2003WO-US017430.
XX
PR   03-JUN-2002; 2002US-0385120P.
PR   04-JUN-2002; 2002US-0385784P.
PR   05-JUN-2002; 2002US-0386041P.
PR   05-JUN-2002; 2002US-0386047P.
PR   06-JUN-2002; 2002US-0386376P.
PR   06-JUN-2002; 2002US-0386453P.
PR   06-JUN-2002; 2002US-0386864P.
PR   06-JUN-2002; 2002US-0387016P.
PR   07-JUN-2002; 2002US-0386796P.
PR   07-JUN-2002; 2002US-0386816P.
PR   07-JUN-2002; 2002US-0386931P.
PR   07-JUN-2002; 2002US-0386942P.
PR   07-JUN-2002; 2002US-0386971P.
PR   07-JUN-2002; 2002US-0387262P.
PR   08-JUN-2002; 2002US-0296960P.
PR   10-JUN-2002; 2002US-0387400P.
PR   10-JUN-2002; 2002US-0387535P.
PR   11-JUN-2002; 2002US-0387610P.
PR   11-JUN-2002; 2002US-0387625P.
PR   11-JUN-2002; 2002US-0387634P.
PR   11-JUN-2002; 2002US-0387668P.
PR   11-JUN-2002; 2002US-0387696P.
PR   11-JUN-2002; 2002US-0387702P.
PR   11-JUN-2002; 2002US-0387836P.
PR   11-JUN-2002; 2002US-0387859P.
PR   12-JUN-2002; 2002US-0387933P.
PR   12-JUN-2002; 2002US-0387934P.
PR   12-JUN-2002; 2002US-0387960P.
PR   12-JUN-2002; 2002US-0388022P.
PR   12-JUN-2002; 2002US-0388096P.
PR   13-JUN-2002; 2002US-0389123P.
PR   14-JUN-2002; 2002US-0389118P.
PR   14-JUN-2002; 2002US-0389120P.

PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.

DR P-PSDB; ADH71638.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 533; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506
```

```
Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346
```

```
Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
             ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
                |||
Db       2784 TGA 2786


RESULT 10
ADH71641
ID   ADH71641 standard; DNA; 2881 BP.
XX
AC   ADH71641;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Human gene of the invention NOV21q SEQ ID NO:537.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
```

KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.

PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71642.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 537; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising

CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 986 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686
```

```
Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy        1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy        1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy        1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy        1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy        1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy        1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
                ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy        1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy        1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy        1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy        2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
                ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2124 AAGGAGCTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy        2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy        2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy        2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy        2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||
Db       2784 TGA 2786
```

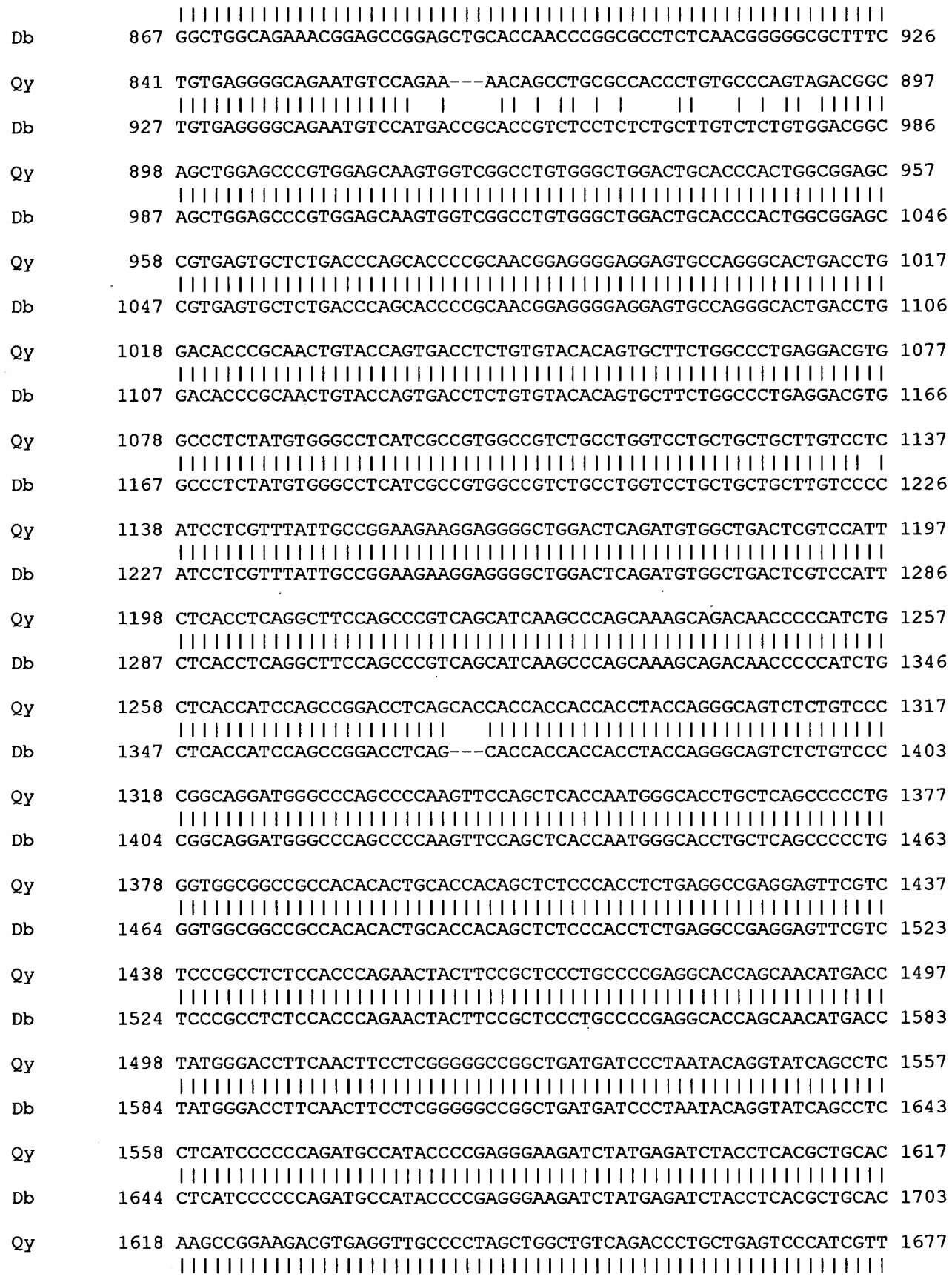
RESULT 11
ADH71629
ID   ADH71629 standard; DNA; 2881 BP.
XX
AC   ADH71629;
XX
DT   25-MAR-2004 (first entry)
XX
DE   Human gene of the invention NOV21k SEQ ID NO:525.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
XX
PF   03-JUN-2003; 2003WO-US017430.
XX
PR   03-JUN-2002; 2002US-0385120P.
PR   04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.

PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71630.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 525; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
```



```
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603
```

```
Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||
Db       2784 TGA 2786
```

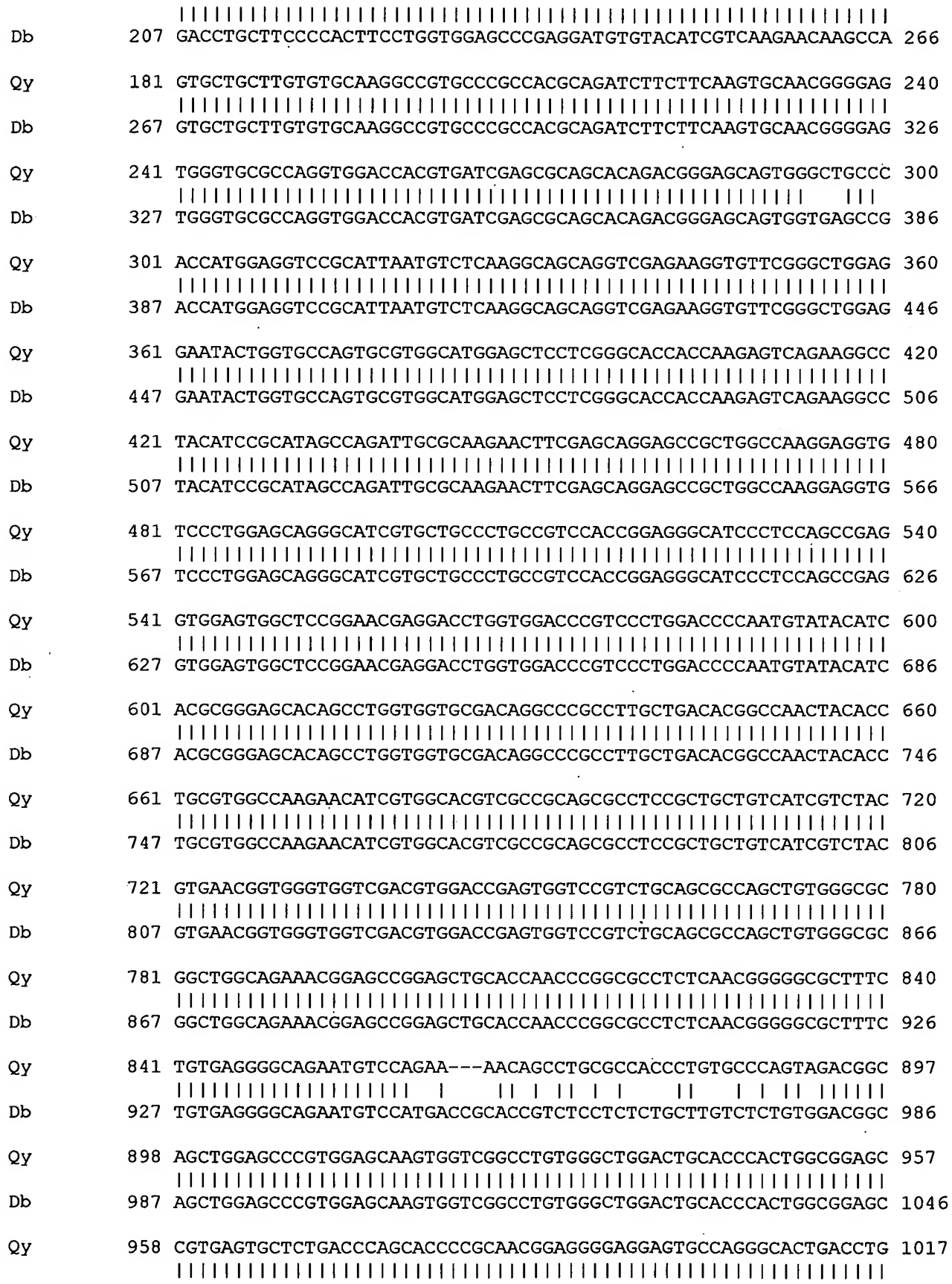
RESULT 12
ADH71631
ID ADH71631 standard; DNA; 2881 BP.
XX
AC ADH71631;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV211 SEQ ID NO:527.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.

PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71632.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 527; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 984 C; 868 G; 503 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
```



```
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGTCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
              ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943
```

```
Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783
```

```
Qy       2695 TGA 2697
              |||
Db       2784 TGA 2786


RESULT 13
ADH71645
ID   ADH71645 standard; DNA; 2881 BP.
XX
AC   ADH71645;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Human gene of the invention NOV21s SEQ ID NO:541.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
XX
PF   03-JUN-2003; 2003WO-US017430.
XX
PR   03-JUN-2002; 2002US-0385120P.
PR   04-JUN-2002; 2002US-0385784P.
PR   05-JUN-2002; 2002US-0386041P.
PR   05-JUN-2002; 2002US-0386047P.
PR   06-JUN-2002; 2002US-0386376P.
PR   06-JUN-2002; 2002US-0386453P.
PR   06-JUN-2002; 2002US-0386864P.
PR   06-JUN-2002; 2002US-0387016P.
PR   07-JUN-2002; 2002US-0386796P.
PR   07-JUN-2002; 2002US-0386816P.
PR   07-JUN-2002; 2002US-0386931P.
PR   07-JUN-2002; 2002US-0386942P.
PR   07-JUN-2002; 2002US-0386971P.
PR   07-JUN-2002; 2002US-0387262P.
PR   08-JUN-2002; 2002US-0296960P.
PR   10-JUN-2002; 2002US-0387400P.
PR   10-JUN-2002; 2002US-0387535P.
PR   11-JUN-2002; 2002US-0387610P.
PR   11-JUN-2002; 2002US-0387625P.
PR   11-JUN-2002; 2002US-0387634P.
PR   11-JUN-2002; 2002US-0387668P.
PR   11-JUN-2002; 2002US-0387696P.
PR   11-JUN-2002; 2002US-0387702P.
PR   11-JUN-2002; 2002US-0387836P.
PR   11-JUN-2002; 2002US-0387859P.
PR   12-JUN-2002; 2002US-0387933P.
PR   12-JUN-2002; 2002US-0387934P.
PR   12-JUN-2002; 2002US-0387960P.
```

PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;

PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71646.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 541; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286
```

```
Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
             |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123
```

```
Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACCGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||
Db       2784 TGA 2786
```

RESULT 14
ADH71627
ID   ADH71627 standard; DNA; 2881 BP.
XX
AC   ADH71627;
XX
DT   25-MAR-2004 (first entry)
XX

DE Human gene of the invention NOV21j SEQ ID NO:523.
XX
KW ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW obesity; diabetes; infectious disease; metabolic syndrome X;
KW dyslipidaemia.
XX
OS Homo sapiens.
XX
PN WO2003102155-A2.
XX
PD 11-DEC-2003.
XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.

PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71628.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 523; 1880pp; English.
XX

CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626
```

```
Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              |||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTGCACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463
```

```
Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1644 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1703

Qy       1618 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1704 AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT 1763

Qy       1678 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1764 AGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT 1823

Qy       1738 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1824 GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC 1883

Qy       1798 TGGG---AGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1854
             ||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1884 TGGGAGCAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAG 1943

Qy       1855 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 1914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1944 CTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGA 2003

Qy       1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCC 2063

Qy       1975 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2064 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTC 2123

Qy       2035 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2124 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183

Qy       2095 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2184 CTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCC 2243

Qy       2155 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2244 CTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAAT 2303

Qy       2215 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2274
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2304 GGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGT 2363

Qy       2275 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2334
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2364 GACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2423

Qy       2335 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2424 AACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2483

Qy       2395 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2454
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
                    |||
Db       2784 TGA 2786
```

```
RESULT 15
ADH71639
ID   ADH71639 standard; DNA; 2881 BP.
XX
AC   ADH71639;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Human gene of the invention NOV21p SEQ ID NO:535.
XX
KW   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW   obesity; diabetes; infectious disease; metabolic syndrome X;
KW   dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO2003102155-A2.
XX
PD   11-DEC-2003.
```

XX
PF 03-JUN-2003; 2003WO-US017430.
XX
PR 03-JUN-2002; 2002US-0385120P.
PR 04-JUN-2002; 2002US-0385784P.
PR 05-JUN-2002; 2002US-0386041P.
PR 05-JUN-2002; 2002US-0386047P.
PR 06-JUN-2002; 2002US-0386376P.
PR 06-JUN-2002; 2002US-0386453P.
PR 06-JUN-2002; 2002US-0386864P.
PR 06-JUN-2002; 2002US-0387016P.
PR 07-JUN-2002; 2002US-0386796P.
PR 07-JUN-2002; 2002US-0386816P.
PR 07-JUN-2002; 2002US-0386931P.
PR 07-JUN-2002; 2002US-0386942P.
PR 07-JUN-2002; 2002US-0386971P.
PR 07-JUN-2002; 2002US-0387262P.
PR 08-JUN-2002; 2002US-0296960P.
PR 10-JUN-2002; 2002US-0387400P.
PR 10-JUN-2002; 2002US-0387535P.
PR 11-JUN-2002; 2002US-0387610P.
PR 11-JUN-2002; 2002US-0387625P.
PR 11-JUN-2002; 2002US-0387634P.
PR 11-JUN-2002; 2002US-0387668P.
PR 11-JUN-2002; 2002US-0387696P.
PR 11-JUN-2002; 2002US-0387702P.
PR 11-JUN-2002; 2002US-0387836P.
PR 11-JUN-2002; 2002US-0387859P.
PR 12-JUN-2002; 2002US-0387933P.
PR 12-JUN-2002; 2002US-0387934P.
PR 12-JUN-2002; 2002US-0387960P.
PR 12-JUN-2002; 2002US-0388022P.
PR 12-JUN-2002; 2002US-0388096P.
PR 13-JUN-2002; 2002US-0389123P.
PR 14-JUN-2002; 2002US-0389118P.
PR 14-JUN-2002; 2002US-0389120P.
PR 14-JUN-2002; 2002US-0389144P.
PR 14-JUN-2002; 2002US-0389146P.
PR 17-JUN-2002; 2002US-0389729P.
PR 17-JUN-2002; 2002US-0389742P.
PR 18-JUN-2002; 2002US-0389884P.
PR 19-JUN-2002; 2002US-0390006P.
PR 19-JUN-2002; 2002US-0390209P.
PR 21-JUN-2002; 2002US-0390763P.
PR 17-JUL-2002; 2002US-0396706P.
PR 06-AUG-2002; 2002US-0401628P.
PR 09-AUG-2002; 2002US-0402156P.
PR 09-AUG-2002; 2002US-0402256P.
PR 09-AUG-2002; 2002US-0402389P.
PR 12-AUG-2002; 2002US-0402786P.
PR 12-AUG-2002; 2002US-0402816P.
PR 12-AUG-2002; 2002US-0402821P.
PR 12-AUG-2002; 2002US-0402832P.
PR 13-AUG-2002; 2002US-0403448P.
PR 13-AUG-2002; 2002US-0403459P.
PR 13-AUG-2002; 2002US-0403531P.
PR 13-AUG-2002; 2002US-0403532P.

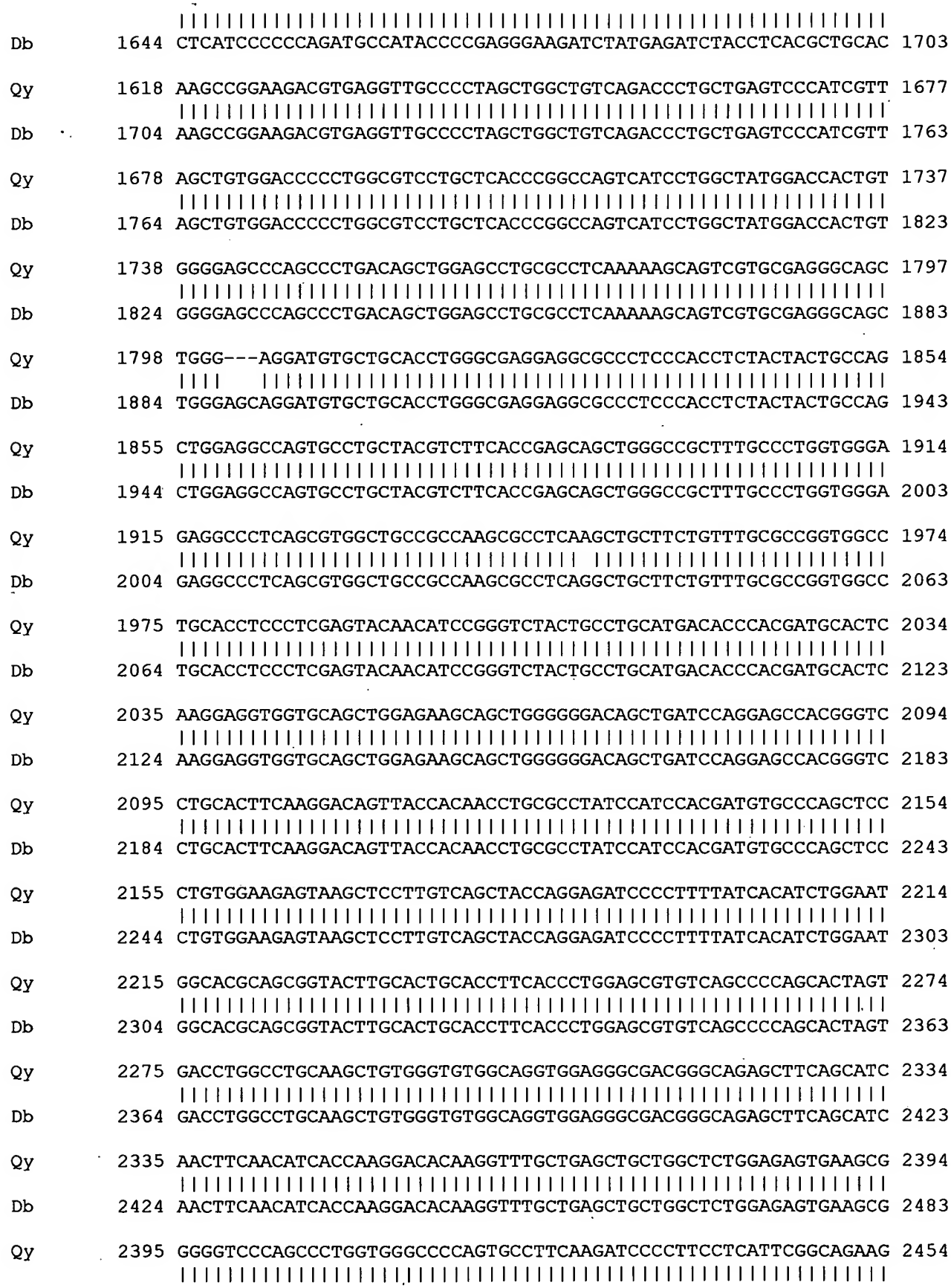
PR 13-AUG-2002; 2002US-0403563P.
PR 13-AUG-2002; 2002US-0406317P.
PR 15-AUG-2002; 2002US-0403617P.
PR 26-AUG-2002; 2002US-0406182P.
PR 26-AUG-2002; 2002US-0406355P.
PR 27-AUG-2002; 2002US-0406240P.
PR 12-SEP-2002; 2002US-0410084P.
PR 20-SEP-2002; 2002US-0412528P.
PR 23-SEP-2002; 2002US-0412731P.
PR 30-SEP-2002; 2002US-0414801P.
PR 30-SEP-2002; 2002US-0414839P.
PR 30-SEP-2002; 2002US-0414840P.
PR 30-SEP-2002; 2002US-0414954P.
PR 09-OCT-2002; 2002US-0417186P.
PR 09-OCT-2002; 2002US-0417406P.
PR 23-OCT-2002; 2002US-0420639P.
PR 28-OCT-2002; 2002US-0421156P.
PR 31-OCT-2002; 2002US-0422690P.
PR 01-NOV-2002; 2002US-0423130P.
PR 05-NOV-2002; 2002US-00423798.
PR 05-NOV-2002; 2002US-0423798P.
PR 12-NOV-2002; 2002US-0425453P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI Zhong H;
XX
DR WPI; 2004-081935/08.
DR P-PSDB; ADH71640.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS Example 21; SEQ ID NO 535; 1880pp; English.
XX
CC The invention relates to a novel isolated polypeptide (NOVX). A
CC polypeptide of the invention has cytostatic, immunomodulator,
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC antilipaemic activity, and may have a use in gene therapy, and as a
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC any of the 303 fully defined nucleotide sequences given in the
CC specification. The polypeptide is useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease. The
CC polypeptide, polynucleotide and antibody are useful in diagnosing,
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence
CC encodes a NOVX polypeptide of the invention.
XX
SQ Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;

Query Match 97.1%; Score 2619.8; DB 12; Length 2881;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

```
Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 146

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 206

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 266

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 326

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        327 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCG 386

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 446

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 506

Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 566

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 626

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 686

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 746

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 806
```

```
Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 866

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 926

Qy        841 TGTGAGGGGCAGAATGTCCAGAA---AACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC 897
              ||||||||||||||||||||  |    || | || | |    ||   | | || ||||||
Db        927 TGTGAGGGGCAGAATGTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGC 986

Qy        898 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC 1046

Qy        958 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTG 1106

Qy       1018 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG 1166

Qy       1078 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC 1226

Qy       1138 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286

Qy       1198 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG 1346

Qy       1258 CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1317
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1347 CTCACCATCCAGCCGGACCTCAG---CACCACCACCACCTACCAGGGCAGTCTCTGTCCC 1403

Qy       1318 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1404 CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG 1463

Qy       1378 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1464 GGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC 1523

Qy       1438 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1524 TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC 1583

Qy       1498 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1584 TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1643

Qy       1558 CTCATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC 1617
```



```
Db       2484 GGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAG 2543

Qy       2455 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2544 ATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAG 2603

Qy       2515 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2604 AAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATG 2663

Qy       2575 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2664 ATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2723

Qy       2635 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2724 GCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGC 2783

Qy       2695 TGA 2697
             |||
Db       2784 TGA 2786
```

Search completed: March 5, 2005, 22:58:58
Job time : 1381 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 19:26:23 ; Search time 424.67 Seconds
(without alignments)
10391.692 Million cell updates/sec

Title: US-10-624-932-1_COPY_46_2742
Perfect score: 2697
Sequence: 1 atggccgtccggcccggcct.....tgtcggaggctgagtgctga 2697

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2252.2	83.5	3014	2	US-08-808-982-1	Sequence 1, Appli
2	2252.2	83.5	3014	3	US-09-306-902A-1	Sequence 1, Appli
3	1552.4	57.6	1787	2	US-08-808-982-2	Sequence 2, Appli
4	1552.4	57.6	1787	3	US-09-306-902A-2	Sequence 2, Appli
5	841.4	31.2	2831	2	US-08-808-982-3	Sequence 3, Appli
6	841.4	31.2	2831	3	US-09-306-902A-3	Sequence 3, Appli
7	833.6	30.9	3008	4	US-09-949-016-4794	Sequence 4794, Ap
8	487	18.1	2736	4	US-09-969-532-9	Sequence 9, Appli
9	487	18.1	3411	4	US-09-969-532-33	Sequence 33, Appl
10	467	17.3	2703	4	US-09-969-532-11	Sequence 11, Appl
11	460.2	17.1	2694	4	US-09-969-532-13	Sequence 13, Appl

	12	439.8	16.3	2661	4	US-09-969-532-15	Sequence 15, Appl
	13	323	12.0	349	4	US-09-471-276-345	Sequence 345, App
	14	293.2	10.9	1968	4	US-09-969-532-31	Sequence 31, Appl
	15	293.2	10.9	2001	4	US-09-969-532-29	Sequence 29, Appl
	16	293.2	10.9	2010	4	US-09-969-532-27	Sequence 27, Appl
	17	293.2	10.9	2043	4	US-09-969-532-25	Sequence 25, Appl
	18	284.8	10.6	1659	4	US-09-969-532-7	Sequence 7, Appli
	19	284.8	10.6	1692	4	US-09-969-532-5	Sequence 5, Appli
	20	284.8	10.6	1701	4	US-09-969-532-3	Sequence 3, Appli
	21	284.8	10.6	1734	4	US-09-969-532-1	Sequence 1, Appli
c	22	269	10.0	771	1	US-08-253-155A-17	Sequence 17, Appl
c	23	181.4	6.7	19326	4	US-09-949-016-16776	Sequence 16776, A
	24	123.4	4.6	114139	4	US-09-949-016-16536	Sequence 16536, A
	25	119	4.4	305	2	US-08-808-982-4	Sequence 4, Appli
	26	119	4.4	305	3	US-09-306-902A-4	Sequence 4, Appli
c	27	80.4	3.0	601	4	US-09-949-016-169999	Sequence 169999,
	28	80.2	3.0	966	4	US-09-969-532-23	Sequence 23, Appl
	29	75.8	2.8	1008	4	US-09-969-532-19	Sequence 19, Appl
c	30	73.8	2.7	601	4	US-09-949-016-170033	Sequence 170033,
c	31	66.4	2.5	601	4	US-09-949-016-170040	Sequence 170040,
c	32	66.4	2.5	601	4	US-09-949-016-170041	Sequence 170041,
	33	63.6	2.4	5784	4	US-09-949-016-462	Sequence 462, App
c	34	63	2.3	601	4	US-09-949-016-170000	Sequence 170000,
c	35	63	2.3	601	4	US-09-949-016-170034	Sequence 170034,
	36	60.2	2.2	2820	4	US-09-854-845-15	Sequence 15, Appl
	37	60.2	2.2	2865	4	US-09-854-845-13	Sequence 13, Appl
	38	60.2	2.2	3105	4	US-09-854-845-5	Sequence 5, Appli
	39	60.2	2.2	3150	4	US-09-854-845-1	Sequence 1, Appli
	40	60.2	2.2	3237	4	US-09-854-845-7	Sequence 7, Appli
	41	60.2	2.2	3282	4	US-09-854-845-3	Sequence 3, Appli
	42	60.2	2.2	3411	4	US-09-854-845-11	Sequence 11, Appl
	43	60.2	2.2	3456	4	US-09-854-845-9	Sequence 9, Appli
	44	60.2	2.2	4074	4	US-09-854-845-17	Sequence 17, Appl
	45	56.4	2.1	5178	4	US-09-949-016-5241	Sequence 5241, Ap

ALIGNMENTS

```
RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
;  GENERAL INFORMATION:
;    APPLICANT:  Tessier-Lavigne, Marc
;    APPLICANT:  Leonardo, E. David
;    APPLICANT:  Hink, Lindsay
;    APPLICANT:  Masu, Masayuki
;    APPLICANT:  Kazuko, Keino-Masu
;    TITLE OF INVENTION:  Netrin Receptors
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  268 BUSH STREET, SUITE 3200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
```

```
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/808,982
;       FILING DATE:
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  OSMAN, RICHARD A
;       REGISTRATION NUMBER:  36,627
;       REFERENCE/DOCKET NUMBER:  UC96-217
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 343-4341
;       TELEFAX:  (415) 343-4342
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3014 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-808-982-1

  Query Match             83.5%;  Score 2252.2;  DB 2;  Length 3014;
  Best Local Similarity   89.7%;  Pred. No. 0;
  Matches 2419;  Conservative    0;  Mismatches  278;  Indels    0;  Gaps    0;

Qy          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||| |||||
Db          1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60

Qy         61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
              || || |||||||||||||||||||||| |||||||| |||||||| ||||||||||
Db         61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120

Qy        121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
              |||||||| |||||||||||||| ||||| ||||| |||||||| ||||||||||||||
Db        121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180

Qy        181 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
              ||| || | |||||||||| ||||| ||||| ||||||||||||||||||| |||||
Db        181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240

Qy        241 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
              ||||| |||||||| || ||||| || || |||||||| ||| | ||||| ||  ||||
Db        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCGACAGCAGCAGCGGATTGCCA 300

Qy        301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
              |||||||||||||| || || || || |||||||||| ||||| ||||| |||||||||
Db        301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360

Qy        361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
              |||||||||||||||| |||||||||||||||||| |||||||| ||||||||||||
Db        361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
```

```
Qy        421 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 480
              |||||||| || |||   |||||||||||||| |||||||||| |||||||||| |||
Db        421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480

Qy        481 TCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 540
              || |||||||| ||||| || || || || || || || || || ||||| ||||| |||
Db        481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540

Qy        541 GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC 600
              |||||||||||  || || |||||||| |||||||| ||||| || |||||||| ||||||
Db        541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600

Qy        601 ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC 660
              |||||||||||||||| || ||||| |||||||||| || ||||||||||||||||||
Db        601 ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC 660

Qy        661 TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC 720
              || |||||||||||||||| || |||||||| ||| |||| || || ||||| || ||
Db        661 TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT 720

Qy        721 GTGAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC 780
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        721 GTGAACGGTGGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT 780

Qy        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC 840
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db        781 GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC 840

Qy        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC 900
              |||||||||||||||||||||||||||||||||||||| |||||||||| || || |||
Db        841 TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC 900

Qy        901 TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT 960
              |||||  ||||||| |||||||| |||||||||| ||||||||||||||||||||||
Db        901 TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC 960

Qy        961 GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC 1020
              ||||||||||||||||||||||||| ||||| |||||||| | |||  ||||||||||
Db        961 GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC 1020

Qy       1021 ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC 1080
              |||||||||||||||||||||||||  | ||||  |||||| |||| ||||||||||
Db       1021 ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT 1080

Qy       1081 CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC 1140
              |||||  | |||||  |||| ||||| || |||||  || ||||| |||| | |||
Db       1081 CTCTACATCGGCCTTGTCGCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA 1140

Qy       1141 CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC 1200
              ||| |||| || || |||||||| |||||||||| |||||||| |||||||||| |||
Db       1141 CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC 1200

Qy       1201 ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC 1260
              ||||| |||||||||||| |||||||||||||||||||||||||||||||| ||||||
Db       1201 ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC 1260
```